THE BIRTH OF THE INFINITE CLUSTER: FINITE-SIZE SCALING IN PERCOLATION

C. Borgs[1], J. T. Chayes[1], H. Kesten[2], and J. Spencer[3]

[1] *Microsoft Research*
One Microsoft Way, Redmond, WA 98052

[2] *Department of Mathematics, Cornell University*
Ithaca, NY 14853

[3] *Courant Institute of Mathematical Sciences*
New York University, 251 Mercer Street, New York, NY 10012

December 2000*

Abstract. We address the question of finite-size scaling in percolation by studying bond percolation in a finite box of side length n, both in two and in higher dimensions. In dimension $d = 2$, we obtain a complete characterization of finite-size scaling. In dimensions $d > 2$, we establish the same results under a set of hypotheses related to so-called scaling and hyperscaling postulates which are widely believed to hold up to $d = 6$.

As a function of the size of the box, we determine the scaling window in which the system behaves critically. We characterize criticality in terms of the scaling of the sizes of the largest clusters in the box: incipient infinite clusters which give rise to the infinite cluster. Within the scaling window, we show that the size of the largest cluster behaves like $n^d\pi_n$, where π_n is the probability at criticality that the origin is connected to the boundary of a box of radius n. We also show that, inside the window, there are typically many clusters of scale $n^d\pi_n$, and hence that "the" incipient infinite cluster is not unique. Below the window, we show that the size of the largest cluster scales like $\xi^d\pi_\xi\log(n/\xi)$, where ξ is the correlation length, and again, there are many clusters of this scale. Above the window, we show that the size of the largest cluster scales like n^dP_∞, where P_∞ is the infinite cluster density, and that there is only one cluster of this scale. Our results are finite-dimensional analogues of results on the dominant component of the Erdős-Rényi mean-field random graph model.

*Revised in May 2001

1. Introduction: Background and Discussion of Results

We dedicate this paper to Joel Lebowitz on the occasion of his 70th birthday. He is an inspiration to us all. We present here the complete version of results announced several years ago in [CPS96] and [Cha98].

Finite-size scaling is the study of corrections to the thermodynamic behavior of an infinite system due to finite-size effects. In particular, this includes the broadening of the transition point into a transition region in a finite system. Here we present an analysis of finite-size scaling for percolation on the hypercubic lattice, both in two and in higher dimensions. Our analysis is based on a number of postulates which are mathematical expressions of the purported scaling behavior in critical percolation in dimensions two through six. We explicitly verify these scaling postulates in two dimension.

We consider bond percolation in a finite subset Λ of the hypercubic lattice $\mathbb{Z}^d$. Nearest-neighbor bonds in Λ are occupied with probability p and vacant with probability $1-p$, independently of each other. Let p_c denote the bond percolation threshold in $\mathbb{Z}^d$, namely the value of p above which there exists an infinite connected cluster of occupied bonds. As a function of the size of the box Λ, we determine the scaling window about p_c in which the system behaves critically. For our purposes, criticality is characterized by the behavior of the distribution of sizes of the largest clusters in the box. We show how these clusters can be identified with the so-called incipient infinite cluster—the cluster of infinite expected size which appears at p_c.

The motivation for this work was threefold: first, to give a finite-dimensional analogue and interpretation of results on the Erdős-Rényi mean-field random graph model; second, to provide rigorous results on finite-size scaling at a continuous transition; and third, to establish detailed results on incipient infinite clusters which correspond closely to results observed by numerical physicists. In this introduction, we will discuss each aspect of the motivation in some detail.

The Random Graph Model

The original motivation for this work was to obtain an analogue of known results on the random graph model of Erdős and Rényi ([ER59], [ER60]; see also [Bol85], [AS92]). The random graph is simply the percolation model on the complete graph, i.e., it is a model on a graph of N sites in which each site is connected to each other site, independently with uniform probability $p(N)$. It turns out that the model has particularly interesting behavior if $p(N)$ scales like $p(N) \approx c/N$ with $c \asymp 1$. Here, as usual, $f \asymp g$ means that there are nonzero, finite strictly positive constants c_1 and c_2, such that $c_1 g \leq f \leq c_2 g$.

Let $W^{(i)}$ denote the random variable representing the size of the i^{th} largest cluster in the system. Erdős and Rényi ([ER59], [ER60]) showed that the model has a *phase transition* at $c=1$ characterized by the behavior of $W^{(1)}$. It turns out that, with probability one,

$$W^{(1)} \asymp \begin{cases} \log N & \text{if} \quad c<1 \\ N^{2/3} & \text{if} \quad c=1 \\ N & \text{if} \quad c>1. \end{cases} \tag{1.1}$$

Moreover, for $c>1$, $W^{(1)}/N$ tends to some constant $\theta(c)>0$, with probability one, while for $c=1$, $W^{(1)}$ has a nontrivial distribution (i.e., $W^{(1)}/N^{2/3} \nrightarrow$ constant) ([ER59], [ER60], [JKLP93], [Ald97]). For $c \leq 1$, the sizes of the second, third, ... , largest clusters are of the same scale as that of the largest cluster, while for $c > 1$ this is not the case: For any fixed $i > 1$, $W^{(i)} \asymp \log N$ for all $c \neq 1$ ([ER59], [ER60]), while at $c = 1$, $W^{(i)} \asymp N^{2/3}$ [Bol84]. The cluster of order N for $c > 1$ is clearly the analogue of the infinite cluster in percolation on finite-dimensional graphs; in the random graph, it is called the *giant component*. As we will see, the clusters of order $\log N$ or smaller are analogues of finite clusters in ordinary percolation. The clusters of order $N^{2/3}$ will turn out to be the analogue of the so-called *incipient infinite cluster* in percolation.

More interestingly, the critical point $c = 1$ is actually broadened into a critical regime by finite-N corrections. It was shown by Bollobás [Bol84] and Łuczak [Łuc90] that the correct parameterization of the critical regime is

$$p(N) = \frac{1}{N}\left(1 + \frac{\lambda_N}{N^{1/3}}\right), \tag{1.2}$$

in the sense that if $\lim_{N\to\infty}|\lambda_N| < \infty$, then $W^{(i)} \asymp N^{2/3}$ for all i; see also the combinatoric tour de force of Janson, Knuth, Łuczak and Pittel [JKLP93] for more detailed properties, including some distributional results on the $W^{(i)}$'s. Finally, it was shown by Aldous that the $W^{(i)}$, rescaled by $N^{2/3}$, have a nontrivial limiting joint distribution which can be calculated from a one-dimensional Brownian motion with time-dependent drift [Ald97].

On the other hand, if $\lim_{N\to\infty}\lambda_N = -\infty$, then $W^{(2)}/W^{(1)} \to 1$ with probability one, whereas if $\lim_{N\to\infty}\lambda_N = +\infty$, then $W^{(2)}/W^{(1)} \to 0$ and $W^{(1)}/N^{2/3} \to \infty$ with probability one. The largest component in the regime with $\lambda_N \to +\infty$ is called the *dominant component*. As we will show, it has an analogue in ordinary percolation.

The initial motivation for our work was to find a finite-dimensional analogue of the above results. To this end, we consider d-dimensional percolation in a box of linear size n, and hence volume $N = n^d$. We ask how the size of the largest cluster in the box behaves as a function of n for $p < p_c$, $p = p_c$ and $p > p_c$. It is straightforward from known results to describe these cluster sizes for *fixed* $p \neq p_c$. However, we are interested mainly in the situation where p varies with n. In particular, we ask whether there is a window about p_c such that the system has a nontrivial cluster size distribution within the window.

Finite-Size Scaling

The considerations of the previous paragraph lead us immediately to the question of *finite-size scaling* (FSS). Phase transitions cannot occur in finite volumes, since all relevant functions are polynomials and thus analytic; nonanalyticities only emerge in the infinite-volume limit. What quantities should we study to see the phase transition emerge as we go to larger and larger volumes?

Before our work, this question had been rigorously addressed in detail only in systems with first-order transitions—transitions at which the correlation length and order parameter are discontinuous ([BoK90], [BI92-1], [BI92-2]). Finite-size scaling at second-order transitions is more subtle due to the fact that the order parameter vanishes at the critical point. For example, in percolation it is believed that the infinite cluster density vanishes at p_c. However, physicists routinely talk about an incipient infinite cluster at p_c. This brings us to our third motivation.

The Incipient Infinite Cluster

At p_c, it is believed that with probability one there is no infinite cluster. On the other hand, the *expected size* of the cluster of the origin is infinite at p_c, see [Ham57], [Kes82], Cor. 5.1, and [AN84]. This suggests that from the perspective of an observer at the origin, all clusters are finite, with larger and larger clusters appearing as one considers larger and larger length scales. Physicists have called the emerging object the incipient infinite cluster.

In the mid-1980's there were two attempts to construct rigorously an object that could be identified as an incipient infinite cluster. Kesten [Kes86] proposed to look at the conditional measure in which the origin is connected to the boundary of a box centered at the origin, by a path of occupied bonds: $P_p^n(\cdot) = P_p(\cdot \mid 0 \leftrightarrow \partial[-n,n]^d)$. Here, as usual, $P_p(\cdot)$ is product measure at bond density p. Observe that, at $p = p_c$, as $n \to \infty$, $P_p^n(\cdot)$ becomes mutually singular with respect to the unconditioned measure $P_p(\cdot)$. Nevertheless, Kesten found that in $d = 2$

$$\lim_{n\to\infty} P_{p_c}^n(\cdot) = \lim_{p\searrow p_c} P_p(\cdot \mid 0 \leftrightarrow \infty). \tag{1.3}$$

Moreover, Kesten studied properties of the infinite object so constructed and found that it has a nontrivial fractal dimension which agrees with the fractal dimension of the physicists' incipient infinite cluster.

Another proposal was made by Chayes, Chayes and Durrett [CCD87]. They modified the standard measure in a different manner than Kesten, replacing the uniform p by an inhomogeneous $p(b)$ which varies with the distance of the bond b from the origin:

$$p(b) = p_c + \frac{\lambda}{1 + \mathrm{dist}(0,b)^\zeta}, \tag{1.4}$$

with λ constant. The idea was to enhance the density just enough to obtain a nontrivial infinite object. In $d = 2$, [CCD87] proved that for $\zeta = 1/\nu$, where ν is the so-called correlation length exponent, the measure $P_{p(b)}$ has some properties reminiscent of the physicists' incipient infinite cluster.

In this work, we propose a third rigorous incipient cluster—namely the largest cluster in a box. This is, in fact, exactly the definition that numerical physicists use in simulations. Moreover, it will turn out to be closely related to the IICs constructed by Kesten and Chayes, Chayes and Durrett. Like the IIC of [Kes86], the largest cluster in a box will have a fractal dimension which agrees with that of the physicists' IIC. Also, our proofs rely heavily on technical estimates from the IIC construction of [Kes86]. More interestingly, the form of the scaling window $p(n)$ for the our problem will turn out to be precisely the form of the enhanced density used to construct the IIC of [CCD87].

Yet a fourth candidate for an incipient infinite cluster is a spanning cluster in a large box, an object studied by Aizenman in [Aiz97]. Let us caution the reader that the terminology in [Aiz97] differs somewhat from ours. While Aizenman reserves the term IIC for an incipient infinite cluster *viewed from a point inside* this cluster (thus implying uniqueness almost by definition), we use the term incipient infinite clusters for the large clusters viewed from the scale of the box under consideration. From this point of view the IIC is not necessarily unique, see below.

Recently, Járai has shown that, viewed from a random point in the IIC, all four notions of the IIC lead to the same distribution on local observables in dimension $d = 2$ [Jár00].

Informal Statement and Heuristic Interpretation of Results

Our results will be stated precisely in Section 3. Here we give an informal statement in terms of the critical exponents of percolation, assuming these exponents exist. Note that our results hold independently of the existence of critical exponents, but they are easier to state informally and to compare to the random graph results (1.1) and (1.2) in terms of these exponents. To this end, let $P_\infty(p)$ denote the infinite cluster density, $\chi^{\text{fin}}(p)$ denote the expected size of finite clusters, $\xi(p)$ denote the correlation length, i.e., the inverse exponential decay rate of the finite cluster connectivity function, and $P_{\geq s}(p)$ denote the probability that the cluster of the origin is of size at least s. Also let $\pi_n(p_c)$ denote the probability at criticality that the origin is connected to the boundary of a hypercube of side $2n$. See Section 2, in particular equations (2.5), (2.15), (2.18), (2.4) and (2.10), for precise definitions. It is believed, but not proved in low dimensions, that the behavior of these quantities as $p \to p_c$ or at $p = p_c$ is described by the following scaling laws:

$$P_\infty(p) \approx |p - p_c|^\beta \quad p > p_c\,, \tag{1.5}$$

$$\chi^{\text{fin}}(p) \approx |p - p_c|^{-\gamma}\,, \tag{1.6}$$

$$\xi(p) \approx |p - p_c|^{-\nu}\,, \tag{1.7}$$

$$P_{\geq s}(p_c) \approx s^{-1/\delta} \tag{1.8}$$

and

$$\pi_n(p_c) \approx n^{-1/\rho}\,. \tag{1.9}$$

In (1.5) - (1.7), $G(p) \approx |p - p_c|^\alpha$ means

$$\lim_{p \to p_c} \frac{\log G(p)}{\log |p - p_c|} = \alpha\,. \tag{1.10}$$

Unless otherwise noted we implicitly assume that the approach is identical from above and below threshold. Similarly, we use $G(n) \approx n^\alpha$ in (1.8) - (1.9) to mean

$$\lim_{n \to \infty} \frac{\log G(n)}{\log n} = \alpha. \tag{1.11}$$

Let Λ_n denote a hypercube of side n and let $W_{\Lambda_n}^{(i)}$ denote the i^{th} largest cluster in this hypercube. Then, under certain "scaling assumptions," we find the asymptotic behavior of $W_{\Lambda_n}^{(1)}$, both for fixed p and, more generally, for p which vary with n. Combining our results at p_c with known results for fixed $p \neq p_c$, we first establish the following analogue of (1.1):

$$W_{\Lambda_n}^{(1)} \asymp \begin{cases} \log n & \text{if} \quad p < p_c \\ n^{d_f} & \text{if} \quad p = p_c \\ n^d & \text{if} \quad p > p_c, \end{cases} \tag{1.12}$$

where we use the suggestive notation

$$d_f = d - 1/\rho \tag{1.13}$$

to indicate that $d - 1/\rho$ is the fractal dimension of our candidate incipient infinite cluster.

Moreover, we show that, under the scaling assumptions, the critical point p_c is broadened into a scaling window of the form

$$p(n) = p_c\left(1 \pm \frac{\lambda}{n^{1/\nu}}\right), \tag{1.14}$$

in the sense that inside the window

$$W^{(1)} \approx n^{d_f}, \quad W^{(2)} \approx n^{d_f}, \quad \cdots, \tag{1.15}$$

while above the window

$$\begin{aligned} &W^{(1)} \approx n^d P_\infty\,, \\ &W^{(1)}/n^{d_f} \to \infty\,, \\ &W^{(2)}/W^{(1)} \to 0, \end{aligned} \tag{1.16 }$$

and below the window

$$W^{(1)}/n^{d_f} \to 0\,, \tag{1.17}$$

where, in fact,

$$W^{(1)} \approx \xi^{d_f} \log(n/\xi)\,. \tag{1.18}$$

The results in (1.14) - (1.18) are established both in expectation and in probability. Note the similarity between the form of the scaling window (1.14) and the bond density (1.4) of the [CCD87] incipient infinite cluster.

Furthermore, within the scaling window, we get results on the distribution of cluster sizes which show that the distribution does not go to a point mass. This is to be contrasted with the behavior above the window, where the normalized cluster size approaches its expectation, with probability one. All of these additional results require some delicate second moment estimates.

Our scaling assumptions, which are described in detail in Section 3, are explicitly proved in dimension $d = 2$, and are believed – but not proved – to hold for d less than the so-called upper critical dimension d_c. The upper critical dimension is the dimension above which the critical exponents assume their Cayley tree values; presumably $d_c = 6$ for percolation.

What would results (1.14) and (1.15) say if we attempted to apply them in the case of random graph model (to which, of course, they do not rigorously apply)? Let us use the widely believed hyperscaling relation $d\nu = \gamma + 2\beta$ and the observation that the volume N of our system is just n^d, to rewrite the window in the form

$$p_n = p_c\left(1 \pm \frac{\lambda}{n^{1/\nu}}\right) = p_c\left(1 \pm \frac{\lambda}{N^{1/d\nu}}\right) = p_c\left(1 \pm \frac{\lambda}{N^{1/(\gamma+2\beta)}}\right). \tag{1.19}$$

Similarly, let us use the hyperscaling relation $d_f/d = \delta/(1+\delta)$ to rewrite the size of the largest cluster as

$$W^{(1)} \approx n^{d_f} \approx N^{d_f/d} \approx N^{\delta/(1+\delta)}. \tag{1.20}$$

Noting that the random graph model is a mean-field model, we expect (and in fact it can be verified [BBCK98]) that $\gamma = 1,\ \beta = 1$ and $\delta = 2$. Using also $p_c = 1/N$, (1.19) suggests a window of the form

$$p(N) = \frac{1}{N}\left(1 \pm \frac{\lambda}{N^{1/3}}\right), \tag{1.21}$$

and within that window

$$W^{(1)} \approx N^{2/3}, \tag{1.22}$$

just the values obtained in the combinatoric calculations on the random graph model. We caution the reader that hyperscaling relations do not apply to the random graph, so that a proper version of the arguments above requires that we deal with a "correlation volume" rather than the correlation length, and that we establish (1.20) directly from the scaling of the cluster size distribution (1.8), rather than by recourse to our finite-dimensional results and a hyperscaling relation. Such arguments can be derived, but are beyond the scope of this paper.

Our results also have implications for finite-size scaling. Indeed, the form of the window tells us precisely how to locate the critical point, i.e., it tells us the correct region about p_c in which to do critical calculations.

Finally, the results tell us that we may use the largest cluster in the box as a candidate for the incipient infinite cluster. Within the window, it is not unique, in the sense that there are many clusters of this scale. However, above the window (even including a region where p is not uniformly greater than p_c as $n \to \infty$), there is a unique cluster of largest scale. This is the analogue of what is called the *dominant component* in the random graph problem.

It is interesting to contrast our results with recent results in high dimensions. As already observed on a heuristic level in [Con85], the validity of hyperscaling is related to the fact that the critical crossing clusters in a box of side length n have size of order $n^{d-1/\rho}$, and that their number is bounded uniformly in n; see [BCKS98] for rigorous results concerning this relationship. Conversely, breakdown of hyperscaling above six dimensions requires, at least on a heuristic level, that at criticality, the number of crossing clusters in a box of side length n grows like n^{d-6}, and that all of them have sizes of order n^4; see again [Con85]. In a similar way, one would expect that the *largest* cluster in a box of side length n is of size n^4, and that there are roughly n^{d-6} clusters of similar size. Indeed, it can be proven [Aiz97] that these results follow from a postulate on the decay of the connectivity function at criticality which is widely believed to hold above six dimensions. Very recently, T. Hara [Har01] used the so-called Lace expansion, in the form developed in [HHS01], to rigorously establish this postulate in sufficiently high dimensions $d \gg 6$.

Methods and Organization

As mentioned above, our results are proved under certain scaling assumptions which we explicitly verify in dimension $d = 2$. Obviously, the results could have been proven directly

– with no assumptions – in $d = 2$, but the resulting proof would have been quite complicated and would not have yielded much insight. Instead, we formulate postulates which we believe characterize critical behavior in all dimensions below the critical dimension d_c, and then prove our results under these postulates. We believe that the postulates are of independent interest since they provide insight into the nature of critical behavior. Indeed, in previous announcements of this work [CPS96] and [Cha98], we used more postulates than we need now. In [BCKS99], we proved that one of these original postulates was implied by several others, in particular that a reasonable assumption on the behavior of crossing probabilities implies certain hyperscaling relations among critical exponents. The proofs in this paper will rely heavily on the results and methods of [BCKS99]. Indeed, [BCKS99] should really be viewed as "Part I" of this paper, since many of our results on the cluster size distribution were derived there. The verification of the postulates in $d = 2$ relies on the constructive two-dimensional methods of [Kes86] and [Kes87].

The organization of this paper is as follows. In Section 2, we give definitions, notations and previous percolation results we will need in our proofs. Our main results are formulated in Section 3. There we first state our postulates, and then state the finite-size scaling results under these postulates. In Section 4, we state many additional results which may be of independent interest, including the results of [BCKS99]. Finally, using these additional results, in Section 5 we prove our main finite-size scaling theorems under the scaling postulates. We believe, but cannot prove, that the scaling postulates should hold up to the upper critical dimension, which is believed to be $d_c = 6$ for percolation. Finally, in Section 6, we prove that the scaling postulates are satisfied in two dimensions. Thus, we have a complete proof of finite-size scaling for percolation in dimension $d = 2$. In Section 7, we give a proof of slightly stronger finite-size scaling results under an alternative set of postulates, and also show that the alternative postulates hold in $d = 2$.

2. Definitions, Notation and Preliminaries

Consider the hypercubic site lattice $\mathbb{Z}^d$, and the corresponding bond lattice $\mathbb{B}_d$ consisting of bonds between all nearest-neighbor pairs in $\mathbb{Z}^d$. Bond percolation on $\mathbb{B}_d$ is defined by choosing each bond of $\mathbb{B}_d$ to be *occupied* with probability p and *vacant* with probability $1-p$, independently of all other bonds. The corresponding product measure on configurations of occupied and vacant bonds is denoted by Pr_p. E_p denotes expectation with respect to the measure Pr_p, and $\mathrm{Cov}_p(\cdot\,;\,\cdot)$ denotes the covariance of two indicator functions with respect to Pr_p: $\mathrm{Cov}_p(A; B) = Pr_p(A \cap B) - Pr_p(A)Pr_p(B)$. A generic configuration is denoted by ω. If $S_1, S_2, S_3 \subset \mathbb{Z}^d$, we say that S_1 is connected to S_2 in S_3, denoted by $\{S_1 \leftrightarrow S_2 \text{ in } S_3\}$, if there exists an occupied path with vertices in S_3 from some site of S_1 to some site of S_2. Maximal connected subsets are called *(occupied) clusters*. The occupied cluster (in the configuration ω) containing the site x is denoted by $\mathcal{C}(x) = \mathcal{C}(x; \omega)$. The size of the cluster $\mathcal{C}$, denoted by $|\mathcal{C}|$, is the number of sites in $\mathcal{C}$. $\mathcal{C}_\infty$ denotes the (unique) infinite cluster, i.e., the occupied cluster with $|\mathcal{C}| = \infty$.

We also consider clusters in a finite box $\Lambda \subset \mathbb{Z}^d$. The connected component of x in $\mathcal{C}(x) \cap \Lambda$ is denoted by $\mathcal{C}_\Lambda(x) = \mathcal{C}_\Lambda(x; \omega)$; this is therefore the collection of all points which are connected to x by an occupied path in Λ. $\mathcal{C}_\Lambda^{(1)}, \mathcal{C}_\Lambda^{(2)}, \cdots \mathcal{C}_\Lambda^{(k)}$ denote the occupied clusters in Λ, ordered from largest to smallest size, with lexicographic order between clusters of the same size. $W_\Lambda^{(i)} = |\mathcal{C}_\Lambda^{(i)}|$ denotes the size of the ith largest cluster in Λ. Finally

$$N_\Lambda(s_1, s_2) = |\{i \mid s_1 \leq W_\Lambda^{(i)} \leq s_2\}| \tag{2.1}$$

denotes the number of clusters in Λ with size between s_1 and s_2, and

$$\widetilde{N}_\Lambda(s_1, s_2) = |\{i \mid s_1 \leq W_\Lambda^{(i)} \leq s_2,\, C_\Lambda^{(i)} \not\leftrightarrow \partial\Lambda\}| \tag{2.2}$$

is the corresponding number of clusters which do not touch the boundary $\partial\Lambda$ of Λ. Here $\partial\Lambda$ is the set of points $x \in \Lambda$ that have distance less than 1 from the complement $\Lambda^c = \mathbb{Z}^d \setminus \Lambda$ of Λ.

Returning now to the model on the full lattice, the cluster size distribution is characterized by

$$P_s = P_s(p) = Pr_p(|\mathcal{C}(\mathbf{0})| = s)\,, \tag{2.3}$$

or alternatively

$$P_{\geq s} = P_{\geq s}(p) = Pr_p(|\mathcal{C}(\mathbf{0})| \geq s)\,. \tag{2.4}$$

The order parameter of the model is the *percolation probability* or infinite-cluster density

$$P_\infty(p) = Pr_p(|\mathcal{C}(\mathbf{0})| = \infty)\,. \tag{2.5}$$

The *critical probability* is

$$p_c = \inf\{p : P_\infty(p) > 0\}\,. \tag{2.6}$$

We consider several connectivity functions: the *(point-to-point) connectivity function*

$$\tau(v, w; p) = Pr_p(v \leftrightarrow w)\,, \tag{2.7}$$

the *finite-cluster (point-to-point) connectivity function*

$$\tau^{\text{fin}}(v, w; p) = Pr_p(v \leftrightarrow w, |\mathcal{C}(v)| < \infty)\,, \tag{2.8}$$

the *point-to-hyperplane connectivity function*

$$\widetilde{\pi}_n(p) = Pr_p\{\exists\ v = (n, \cdot)\,\text{such that}\,\mathbf{0} \leftrightarrow v\} \tag{2.9}$$

($v = (n, \cdot)$ means that the first coordinate of v equals n), and the *point-to-box connectivity function*

$$\pi_n(p) = Pr_p\{\mathbf{0} \leftrightarrow \partial B_n(\mathbf{0})\}\,, \tag{2.10}$$

where

$$B_n(v) = \{w \in \mathbb{Z}^d : |v - w|_\infty \leq n\} = [-n, n]^d \cap \mathbb{Z}^d\,, \tag{2.11}$$

with $|\cdot|_\infty$ denoting the ℓ_∞-norm. Notice that $\pi_n(p)$ and $\widetilde{\pi}_n(p)$ are equivalent, in the sense that

$$\widetilde{\pi}_n(p) \leq \pi_n(p) \leq 2d\widetilde{\pi}_n(p)\,. \tag{2.12}$$

A quantity which for $p > p_c$ behaves much like $\tau^{\text{fin}}(x, y; p)$ is the covariance:

$$\tau^{\text{cov}}(v, w; p) = \text{Cov}_p(v \leftrightarrow \infty; w \leftrightarrow \infty) \tag{2.13}$$

(see [CCGKS89], Section 6). We also consider several susceptibilities:

$$\chi(p) = E_p(|\mathcal{C}(\mathbf{0})|) = \sum_v \tau(\mathbf{0}, v; p)\,, \tag{2.14}$$

$$\chi^{\text{fin}}(p) = E_p(|\mathcal{C}(\mathbf{0})|, |\mathcal{C}(\mathbf{0})| < \infty) = \sum_v \tau^{\text{fin}}(\mathbf{0}, v; p) = \sum_{s<\infty} sP_s(p) \tag{2.15}$$

and

$$\chi^{\text{cov}}(p) = \sum_v \tau^{\text{cov}}(\mathbf{0}, v; p)\,. \tag{2.16}$$

Finally, we introduce the quantity

$$s(n) = (2n)^d\, \pi_n(p_c)\,. \tag{2.17}$$

As we will see, $s(n)$ is the order of magnitude of the size of the largest critical clusters on scale n.

Length scales in the model are naturally expressed in terms of the correlation length $\xi(p)$, defined by the limit

$$1/\xi(p) = -\lim_{|v|_\infty \to \infty} \frac{1}{|v|_\infty} \log \tau^{\text{fin}}(\mathbf{0}, v; p) \tag{2.18}$$

taken with v along a coordinate axis. We will use the fact that $\xi(p) < \infty$ for all $p \neq p_c$ and $\xi(p) \to \infty$ as $p \uparrow p_c$ (see Grimmett [Gri99], Theorem 6.49 and equation (6.57) for $p < p_c$; for $p > p_c$ this follows from Grimmett and Marstrand [GM90]). While it is also believed that $\xi(p) \to \infty$ as $p \downarrow p_c$, this is rigorously known only for $d = 2$.

Alternatively, lengths may be expressed in terms of the finite-size scaling correlation length $L_0(p, \varepsilon)$, introduced in [CCF85] and studied in [CCF85], [CCFS86] and [Kes87]. For $p < p_c$, $L_0(p, \varepsilon)$ is defined in terms of the crossing probabilities of rectangles, the so-called *sponge crossing probabilities*:

$$\begin{aligned} R_{L,M}(p) = Pr_p\{ &\exists \text{ occupied bond crossing of } [0, L] \times [0, M] \cdots \times [0, M] \\ &\text{in the 1-direction}\}\,. \end{aligned} \tag{2.19}$$

Observing that, for $p < p_c$, the sponge crossing probability $R_{L,3L}(p) \to 0$ as $L \to \infty$, we define

$$L_0(p) = L_0(p, \varepsilon) = \min\{L \geq 1 \mid R_{L,3L}(p) \leq \varepsilon\} \quad \text{if } p < p_c\,. \tag{2.20}$$

Using the methods and results of [ACCFR83], [CC86], [CCF85] and [Kes87], it is straightforward to show that there exists $a(d) > 0$ such that for $\varepsilon < a(d)$, the scaling behavior of $L_0(p, \varepsilon)$ is independent of ε for $p < p_c$, in the sense that $L_0(p, \varepsilon_1)/L_0(p, \varepsilon_2)$ is bounded away from 0 and infinity for two fixed values $\varepsilon_1, \varepsilon_2 < a(d)$. This scaling behavior is also essentially the same as that of the standard correlation length $\xi(p)$. More specifically, for $0 < \varepsilon < a(d)$, there exist constants $c_1 = c_1(d)$, $c_2 = c_2(d, \varepsilon) < \infty$ such that[1]

$$\frac{1}{L_0(p,\varepsilon)} \le \frac{1}{\xi(p)} \le \frac{c_1 \log L_0(p,\varepsilon) + c_2}{L_0(p,\varepsilon) - 1}, \quad p < p_c \,. \tag{2.21}$$

Hereafter we will assume that $\varepsilon < a(d)$; we usually suppress the ε-dependence in our notation.

For $p > p_c$, it is natural to define $L_0(p, \varepsilon)$ in terms of a suitable finite-cluster analogue of the sponge-crossing probability $R_{L,M}(p)$, see [CC87], eq. (53). For technical reasons, it is convenient, however, to consider instead crossings in an annulus

$$H_{L,M} = \mathbb{Z}^d \cap [-L, L+M]^d \setminus (0, M)^d \,, \tag{2.22}$$

with inner and outer boundaries $\partial_I H_{L,M}$ and $\partial_E H_{L,M}$. We say that an occupied cluster $\mathcal{C}_H$ in $H = H_{L,M}$ is *H-finite* if $H \setminus \mathcal{C}_H$ contains a path – occupied or not – that connects $\partial_I H$ to $\partial_E H$. Let

$$\begin{aligned} S^{\text{fin}}_{L,M}(p) = Pr_p\{ &\exists \text{ an occupied } H\text{-finite cluster } \mathcal{C}_H \text{ in } H = H_{L,M} \\ &\text{that connects } \partial_I H \text{ to } \partial_E H\} \,, \end{aligned} \tag{2.23}$$

with the convention $S^{\text{fin}}_{0,M}(p) = 1$. We define

$$L_0(p) = L_0(p,\varepsilon) = 1 + \max\{L \ge 0 : S^{\text{fin}}_{L,L}(p) \ge \varepsilon\} \qquad \text{if} \quad p > p_c \,, \tag{2.24}$$

and more generally, for $x \ge 1$,

$$L_0(p,\varepsilon;x) = 1 + \max\{L \ge 0 : S^{\text{fin}}_{L,\lfloor xL \rfloor}(p) \ge \varepsilon\} \qquad \text{if} \quad p > p_c \,. \tag{2.25}$$

Note that $L_0(p, \varepsilon; x)$ may be finite or infinite, depending on whether or not there exists an $L_0 < \infty$ such that $S^{\text{fin}}_{L,\lfloor xL \rfloor}(p) < \varepsilon$ for all $L \ge L_0$. We expect that this definition coincides, say in the sense of equation (2.21) (with an $x-$dependent constant c_2, and $c_1(d) = 0$), with the standard correlation length $\xi(p)$ above threshold. However, we are not able to prove this in $d \ge 3$, since the rescaling techniques of [ACCFR83] do not work for finite-cluster crossings. In $d = 2$, we can use a Harris ring construction [Har60] in conjunction with the Russo-Seymour-Welsh Lemma ([Rus78], [SW78]) to show that this definition is equivalent to $\xi(p)$; see Section 6.

An important quantity in the high-density phase is the surface tension $\sigma(p)$; see [AC-CFR83] for the precise definition. By analogy with the definition of a finite-size scaling correlation length below threshold, we define a finite-size scaling inverse surface tension as

$$A_0(p) = A_0(p,\varepsilon) = \min\{L^{d-1} \ge 1 \mid R_{L,3L}(p) \ge 1 - \varepsilon\} \quad \text{if } p > p_c \,. \tag{2.26}$$

[1] K. Alexander [Ale96] has shown that one can take $c_1(d = 2) = 0$ in (2.21)

It is easy to see that $A_0(p)$ is well-defined and finite for all $p > p_c$. Indeed, $p > p_c$ implies $P_\infty(p) > 0$, which in turn implies that the probability of the event $|\mathcal{C}(x)| < \infty$ for all $x \in \mathbb{Z}^d \cap [0, L]^d$ goes to zero as $L \to \infty$. Since this probability is bounded from below by $(1 - R_{L,3L}(p))^{2d}$ (cf. the proof of Lemma 4.4), this implies that $R_{L,3L}(p) \to 1$ as $L \to \infty$, and hence $A_0(p)$ is well-defined and finite. We expect that $A_0(p)$ is equivalent to the inverse surface tension[2] $1/\sigma(p)$, which in turn should be equivalent to $\xi^{d-1}(p)$ below the critical dimension d_c (presumably $d_c = 6$). Again, we are only able to prove this equivalence in $d = 2$.

While the behavior of $L_0(p)$ below p_c is well understood in general dimension, much less is know about $L_0(p)$ or $A_0(p)$ above p_c. In particular, below p_c, it is easy to see that $L_0(p)$ is monotone increasing, left continuous and piecewise constant. Moreover,

$$L_0(p) \uparrow \infty \quad \text{as} \quad p \uparrow p_c\,, \tag{2.27}$$

because $R_{L,3L}(p_c)$ is bounded away from 0 (e.g., by Theorem 5.1 in [Kes82]). Furthermore, the jumps in $L_0(p)$ are uniformly bounded on a logarithmic scale. In particular, by the methods of [ACCFR83], [CC86], [CCF85] and [Kes87], we have

$$R_{2L,6L} \leq \frac{1}{a(d)} R_{L,3L}^2\,, \tag{2.28}$$

which in turn implies

$$\lim_{\delta \to 0} \frac{L_0(p+\delta)}{L_0(p)} \leq 2\,, \tag{2.29}$$

provided $p < p_c$ and $\varepsilon < a(d)$. By contrast, none of these properties are known for $L_0(p)$ above p_c. Next consider $A_0(p)$, which, almost by definition, is monotone decreasing and right continuous. However, in general dimension, we do not have a proof that $A_0(p)$ diverges as $p \downarrow p_c$, nor do we have a bound of the form (2.29). We will therefore require several postulates on the behavior of $L_0(p)$ and $A_0(p)$ above p_c.

3. Statement of Postulates and Theorems

3.1. The Scaling Postulates.

Most of our theorems are established under a set of assumptions which we can verify explicitly in two dimensions, and which we expect to be true for all dimensions not exceeding the critical dimension d_c (presumably $d_c = 6$). We call these assumptions the *Scaling Postulates*, since they follow from the type of scaling typically assumed in the physics literature. Since $L_0(p)$ and $A_0(p)$ depend on ε, see equations (2.20), (2.24) and (2.26), many of our postulates implicitly involve the constant ε. We assume that they are true for all nonzero $\varepsilon < \varepsilon_0$, where $\varepsilon_0 = \varepsilon_0(d)$ is a suitable constant. We write our postulates in terms of the equivalence symbol $\asymp$. Here

$$F(p) \asymp G(p) \tag{3.1}$$

[2] Using Proposition 3 of [CC87], one can actually prove that $A_0(p) \leq \text{const}\,\sigma(p)^{-1}$ for all $d \geq 2$. We do not expect that the opposite inequality holds for $d >$ the critical dimension, d_c, since such an inequality — together with the usual assumption that $\sigma(p) \to 0$ as $p \downarrow p_c$ — would imply that $A_0(p) \to \infty$ as $p \downarrow p_c$ for $d > d_c$, which is believed to be false, see Section 3.3.

means that there are lower and upper bounds of the form

$$C_1 F(p) \le G(p) \le C_2 F(p) \tag{3.2}$$

where $C_1 > 0$ and $C_2 < \infty$ are constants which do not depend on p, as long as p is uniformly bounded away from zero or one, but which may depend on the constants ε, $\widetilde{\varepsilon}$ or x appearing explicitly or implicitly in the postulates. Occasionally, p is further restricted to lie on one side of p_c. Similarly $F(n) \asymp G(n)$ means that

$$C_1 F(n) \le G(n) \le C_2 F(n)$$

for some constants $0 < C_1 \le C_2 < \infty$.

Our scaling postulates are

(I) $L_0(p) \to \infty$ as $p \downarrow p_c$;

(II) $A_0(p) \asymp L_0^{d-1}(p) \asymp L_0^{d-1}(p, \widetilde{\varepsilon}; x)$ provided $p > p_c$, $x \ge 1$ and $0 < \widetilde{\varepsilon} < \varepsilon_0$;

(III) There are constants $D_1 > 0$ and $D_2 < \infty$ such that

$$D_1 \le \frac{\pi_n(p)}{\pi_n(p_c)} \le D_2 \quad \text{if} \quad n \le L_0(p);$$

(IV) There are constants $D_3 > 0$ and $\rho_1 > \frac{2}{d}$ such that

$$\frac{\pi_m(p_c)}{\pi_n(p_c)} \ge D_3 \Big(\frac{m}{n}\Big)^{-1/\rho_1} \quad \text{if} \quad m \ge n \ge 1;$$

(V) There is a constant D_4 such that

$$\chi^{\text{cov}}(p) \le D_4 L_0^d(p) \pi^2_{L_0(p)}(p_c) \quad \text{and} \quad \chi^{\text{fin}}(p) \le D_4 L_0^d(p) \pi^2_{L_0(p)}(p_c)$$

if $p > p_c$;

(VI) $\pi_{L_0(p)}(p_c) \asymp P_\infty(p)$ if $p > p_c$;

(VII) There are constants $D_5, D_6 < \infty$ such that

$$P_{\ge ks(L_0(p))}(p) \ge D_5 e^{-D_6 k} P_{\ge s(L_0(p))}(p) \quad \text{if} \quad p < p_c \quad \text{and} \quad k \ge 1.$$

We shall have some comments on the interpretation of the postulates and other remarks after we state our theorems.

3.2. Statement of the Main Results.

A central concept in our theorems is the notion of a scaling window in which the system behaves critically. This can best be described by the function

$$g(p,n) := \begin{cases} -\frac{n}{L_0(p)} & \text{if } p < p_c \\ 0 & \text{if } p = p_c \\ \frac{n}{L_0(p)} & \text{if } p > p_c. \end{cases} \tag{3.3}$$

It will be seen that a sequence of systems with density p_n behaves critically – as far as size of large clusters is concerned – in the finite boxes

$$\Lambda_n := \{v \in \mathbb{Z}^d \mid -n \le v_i < n, i = 1, \dots, d\} \tag{3.4}$$

if

$$p_n \to p \text{ and } \limsup_{n\to\infty} |g(p_n, n)| < \infty. \tag{3.5}$$

If this is the case we shall say that the (sequence of) systems are *inside the scaling window*. We shall say that the systems are *below* (respectively *above*) *the scaling window* if $g(p_n, n) \to -\infty$ (respectively, $g(p_n, n) \to \infty$). These regimes correspond to subcritical, respectively supercritical behavior. In particular we must have $p_n < p_c$ eventually if $\{p_n\}$ lies below the scaling window, and $p_n > p_c$ eventually if $\{p_n\}$ lies above the scaling window. Our theorems below give many details of the finite-size scaling behavior of the system inside, above, and below the scaling window. They confirm the folklore that within distances of the order of the correlation length the system behaves critically. Specifically, we make this statement precise for the behavior of the size of the large clusters. Unfortunately we cannot derive this from the definition of correlation length only. One of our basic assumptions is that within the correlation length the point-to-box connectivity behaves as it does at the critical point (see Postulate III).

In order to state these theorems, we again use the symbol $\asymp$, this time for two sequences a_n and b_n of real numbers . We write

$$a_n \asymp b_n \tag{3.6}$$

if

$$0 < \liminf_{n\to\infty} \frac{a_n}{b_n} \le \limsup_{n\to\infty} \frac{a_n}{b_n} < \infty\,. \tag{3.7}$$

$|\Lambda_n|$ denotes the number of sites in Λ_n; thus $|\Lambda_n| = (2n)^d$. We remind the reader that Postulates (I)–(VII) are verified for $d = 2$ in Section 6. Thus all the conclusions of our theorems hold in the two-dimensional case.

Our first theorem characterizes the scaling window in terms of the *expectation* of the largest cluster sizes.

Theorem 3.1.

i) Suppose that Postulates (I)–(IV) hold. If $\{p_n\}$ *is inside the scaling window, i.e., if* $\limsup_{n\to\infty} |g(p_n, n)| < \infty$*, and* $i \in \mathbb{N}$*, then*

$$E_{p_n}\{W^{(i)}_{\Lambda_n}\} \asymp s(n)\,. \tag{3.8}$$

ii) Suppose that Postulates (I)–(IV) and (VII) hold. If $\{p_n\}$ *is below the scaling window, i.e.,* $g(p_n, n) \to -\infty$*, then*

$$E_{p_n}\{W^{(1)}_{\Lambda_n}\} \asymp s(L_0(p_n)) \log \frac{n}{L_0(p_n)}\,. \tag{3.9}$$

iii) Suppose that Postulates (II), (V) and (VI) hold. If $\{p_n\}$ *is above the scaling window, i.e.,* $g(p_n, n) \to \infty$*, then*

$$\frac{E_{p_n}\{W^{(1)}_{\Lambda_n}\}}{|\Lambda_n| P_\infty(p_n)} \to 1 \qquad \textit{as} \qquad n \to \infty\,, \tag{3.10}$$

and

$$\frac{E_{p_n}\{W^{(2)}_{\Lambda_n}\}}{|\Lambda_n| P_\infty(p_n)} \to 0 \qquad \textit{as} \qquad n \to \infty\,. \tag{3.11}$$

The next theorem tells us about the *distribution* of the largest cluster sizes above the scaling window.

Theorem 3.2. *Suppose that Postulates (II), (V) and (VI) hold. Let* $\{p_n\}$ *be above the scaling window. Then, as* $n \to \infty$*,*

$$\frac{W^{(1)}_{\Lambda_n}}{|\Lambda_n| P_\infty(p_n)} \to 1 \qquad \textit{in probability}\,. \tag{3.12}$$

The next theorem gives information about the distribution of the cluster sizes inside the scaling window. It shows that, in this regime, the tails of the distribution of $W^{(i)}_{\Lambda_n}/E\{W^{(1)}_{\Lambda_n}\}$ decay, but the distribution does not go to a delta function. This should be contrasted with the behavior (3.12), which shows that above the scaling window the distribution of $W^{(1)}_{\Lambda_n}/E\{W^{(1)}_{\Lambda_n}\}$ does tend to a delta function.

Theorem 3.3. *Suppose that Postulates (I)–(IV) hold. Let* $\{p_n\}$ *lie inside the scaling window.*

i) For all $i < \infty$*,*

$$\liminf_{n\to\infty} Pr_{p_n}\Big\{K^{-1} \le \frac{W^{(i)}_{\Lambda_n}}{E_{p_n}\{W^{(i)}_{\Lambda_n}\}} \le K\Big\} \to 1 \quad \textit{as} \quad K \to \infty\,. \tag{3.13}$$

ii) For each $K < \infty$ and all $i < \infty$,

$$\limsup_{n\to\infty} Pr_{p_n}\Big\{\frac{W^{(i)}_{\Lambda_n}}{E_{p_n}\{W^{(i)}_{\Lambda_n}\}} \geq K^{-1}\Big\} < 1\,. \tag{3.14}$$

We have one more theorem for p inside the scaling window. This concerns the number of clusters on scales $m < n$. Before stating the theorem, we point out that, due to (3.8), the "incipient infinite cluster" inside the scaling window is not unique, in the sense that $W^{(2)}_{\Lambda_n}$ is of the same scale as $W^{(1)}_{\Lambda_n}$. This should be contrasted with the behavior of $W^{(2)}_{\Lambda_n}/W^{(1)}_{\Lambda_n}$ above the scaling window (see (3.10) and (3.11)), a remnant of the uniqueness of the infinite cluster above p_c. The next theorem relates the non-uniqueness of the "incipient infinite cluster" inside the scaling window to the property of scale invariance at p_c. We remind the reader that the quantities N_{Λ_n} and $\widetilde{N}_{\Lambda_n}$ are defined in equation (2.1) and (2.2).

Theorem 3.4. *Suppose that Postulates (I)–(IV) hold. Let $\{p_n\}$ lie inside the scaling window. Then there exist strictly positive, finite constants σ_1, σ_2, C_1 and C_2 (all depending on the sequence $\{p_n\}$, but not on n, m or k) such that*

$$C_1\Big(\frac{n}{m}\Big)^d \leq E_{p_n}\big\{\widetilde{N}_{\Lambda_n}(s(m), s(km))\big\} \leq E_{p_n}\big\{N_{\Lambda_n}(s(m), s(km))\big\} \leq C_2\Big(\frac{n}{m}\Big)^d\,, \tag{3.15}$$

provided m and k are strictly positive integers with $k \geq \sigma_1$ and $\sigma_2 m \leq n$.

Our next theorem gives the behavior of the $W^{(i)}_{\Lambda_n}$ when p is below the scaling window.

Theorem 3.5. *Suppose that Postulates (I)–(IV) and (VII) hold. Let $\{p_n\}$ lie below the scaling window. Then, for each fixed i,*

$$\liminf_{n\to\infty} Pr_{p_n}\Big\{K^{-1} \leq \frac{W^{(i)}_{\Lambda_n}}{s(L_0(p_n))\log\frac{n}{L_0(p_n)}} \leq K\Big\} \to 1 \text{ as } K\to\infty. \tag{3.16}$$

As mentioned before, we expect the Scaling Postulates to hold for all $d \leq d_c = 6$. The next theorem states that they do hold if $d = 2$.

Theorem 3.6. *The Postulates (I) — (VII) hold in $d = 2$.*

Notice that in Theorem 3.3 ii) (in conjunction with (3.8)), we prove that inside the scaling window the support of $W^{(i)}_{\Lambda_n}/s(n)$ is not bounded away from 0. We would expect that this support is also unbounded above and that this should be easy to prove from Postulate (VII), which states in a way that the support of $|\mathcal{C}(\mathbf{0})|/s(L_0(p))$ is unbounded. However we have been unable to derive this from the Postulate (VII). Instead, in Section 7, we consider an alternative postulate, Postulate (VII alt), which says roughly that clusters of size of order $s(L_0(p))$ and distance of order $L_0(p)$ have a reasonable chance of being connected to each other. In that section, we prove the following theorem.

Theorem 3.7.
i) Suppose Postulates (I) – (IV) and (VII alt) hold. Let $\{p_n\}$ be inside the scaling window and let $i \in \mathbb{N}$. Then

$$\limsup_{n\to\infty} Pr_{p_n}\Big\{\frac{W^{(i)}_{\Lambda_n}}{E_{p_n}\{W^{(i)}_{\Lambda_n}\}} \leq K\Big\} < 1 \qquad \textit{for all } K < \infty.$$

ii) Postulate (VII alt) holds in $d = 2$.

3.3 Comments on the Postulates and Further Remarks.

The interpretation of our postulates is as follows. The first tells us that the approach to p_c is critical – i.e., continuous or second-order – from above p_c. The second postulate is the assumption of equivalence of length scales above p_c: namely, Widom scaling, dimensionally relating the surface tension to the correlation length, together with the equivalence of the finite-size scaling lengths at various values of $x \geq 1$ and $\varepsilon \in (0, \varepsilon_0)$. This postulate is not expected to hold above the critical dimension. In fact, it is not even believed that $A_0(p) \to \infty$ as $p \downarrow p_c$, because this would imply that the crossing probability $R_{L,3L}(p_c)$ is bounded away from 1 uniformly in L. But uniform boundedness of crossing probabilities implies hyperscaling [BCKS99], which is not believed to hold above the upper critical dimension d_c. Postulate (III) tell us that the system within the correlation length behaves as it does at threshold, at least as characterized by the behavior of the point-to-box connectivity function. Postulate (IV) implies that the connectivity function has a lower bound of power law behavior at threshold. Especially Postulates (III) and (IV) turn out to imply more than is immediately apparent. Proposition 4.6 states that the cluster size distribution for clusters with diameters up to the correlation length behaves like the corresponding distribution at threshold. This proposition also gives us a hyperscaling relation between the exponents δ and ρ, assuming that these exponents exist. We also obtain a scaling relation for $\chi(p)$ in Proposition 4.8. Assuming power laws for χ and L_0, and the relation (4.24), the assumed bound on ρ_1 in Postulate (IV) is equivalent to the very weak bound $\gamma > 0$. But it is known ([AN84]) that $\chi(p) \geq C_1(p_c - p)^{-1}, p < p_c$, i.e., $\gamma \geq 1$ if it exists. In the light of this, Postulate (IV) seems very reasonable. The fifth and sixth postulates give various exponent relations, again provided that these exponents exist. Finally, the last postulate states that (in the subcritical region) $s(L_0(p))$ is the natural scale for the cluster size distribution and that on this scale the tail of the distribution does not decay faster than exponentially. Proposition 4.8 provides an inequality in the opposite direction, i.e., this decay is at least exponentially fast. See also Remark vi) below.

Remarks. i) Assuming the existence of the exponent ρ, see (1.9), Theorem 3.1 implies that inside the scaling window the largest, second largest, third largest,..., clusters scale like n^{d_f}, with $d_f = d - 1/\rho$, while below the scaling window the size of the largest cluster (and hence of all clusters) goes to zero on the scale n^{d_f}.

ii) By Postulate (VI), and Lemma 4.5 below,

$$\frac{|\Lambda_n| P_\infty(p_n)}{s(n)} = \frac{P_\infty(p_n)}{\pi_n(p_c)} \asymp \frac{\pi_{L_0(p_n)}(p_c)}{\pi_n(p_c)} \to \infty \tag{3.17}$$

above the scaling window. Statement iii of Theorem 3.1 therefore implies that

$$\frac{E_{p_n}\{W^{(1)}_{\Lambda_n}\}}{s(n)} \to \infty \qquad \text{as} \qquad n \to \infty \tag{3.18}$$

above the scaling window.

iii) Assume that the critical exponent ν, see equation (1.7), exists, and that an equivalence of the form (2.21) holds for $p > p_c$ as well. Choose $p_n^- = \sup\{p < p_c : L_0(p) \le n\}$. Then by (2.29), $L_0(p_n^-) \asymp n$. Moreover,

$$L_0(p_n^-) \approx \xi(p_n^-) \approx |p_n^- - p_c|^{-\nu} \tag{3.19}$$

so that $p_c - p_n^- \approx n^{-1/\nu}$. Finally, $\{p_n\}$ is below the scaling window if $\liminf_{n\to\infty} \log(p_c - p_n)/\log n > -1/\nu$. Similar statements hold to the right of p_c with $p_n^+ := \inf\{p > p_c : L_0(p) \le n\}$, provided we make the further assumption that

$$\limsup_{p\downarrow p_c} \lim_{\delta\downarrow 0} \frac{L_0(p-\delta)}{L_0(p)} < \infty.$$

Thus under these various assumptions the scaling window has width $n^{-1/\nu}$.

It should be pointed out, though, that at present we do not have enough rigorous knowledge of the behavior of $L_0(p)$ as a function of p to define the scaling window in terms of the behavior of $(p_n - p_c)/g_n^\pm$ for suitable sequences $\{g_n^\pm\}$. For instance, it is not known that there exists a sequence $\{g_n^-\}$ of positive numbers such that $n/L_0(p_n) \to \infty$ is equivalent to $(p_c - p_n)/g_n^- \to \infty$ for $p_n < p_c$.

iv) It follows from (3.11) and Markov's inequality that

$$\frac{W^{(2)}_{\Lambda_n}}{|\Lambda_n| P_\infty(p_n)} \to 0 \qquad \text{in probability} \tag{3.20}$$

above the scaling window. Combined with (3.12) this implies that, as $n \to \infty$,

$$\frac{W^{(2)}_{\Lambda_n}}{W^{(1)}_{\Lambda_n}} \to 0 \qquad \text{in probability}\,, \tag{3.21}$$

provided $g(p_n, n) \to \infty$.

v) In a similar way, it follows from (3.9) that, as $n \to \infty$,

$$\frac{W^{(1)}_{\Lambda_n}}{s(n)} \to 0 \qquad \text{in probability}\,, \tag{3.22}$$

provided $g(p_n, n) \to -\infty$.

4. Auxiliary results

In this section, which is split into two subsections, we state several useful auxiliary results, most of which have been already proved in [BCKS99], which we will need for our proofs in Section 5. The first subsection gives a fundamental moment estimate and an exponential tail estimate for cluster sizes. These estimates show a close relationship between the diameter and the size or volume of a large cluster. A cluster in Λ_n of diameter small with respect to n usually has a volume which is small with respect to $s(n)$. We believe – but could not prove – that the converse also holds, namely that a cluster in Λ_n of diameter of order n has with high probability a volume bigger than a small multiple of $s(n)$. The second subsection contains various important properties of the quantities $\pi_n, P_s, P_{\geq s}$ and χ which are akin to the postulates.

Throughout, the basic parameter p is bounded away from 0 and 1, that is we restrict p to $\zeta_0 \leq p \leq 1 - \zeta_0$ for some small strictly positive ζ_0. No further mention of ζ_0 will be made. Many constants C_i appear in this paper. These are always *finite and strictly positive*, even when this is not indicated. In different formulae the same symbol C_i may denote different constants. All these constants depend on ε, d, ζ_0 and the constants which appear in the postulates. This dependence will not be indicated in the notation. $I[A]$ denotes the indicator function of the event A.

All results in this section are proven under Postulates (I) — (IV) or a subset of these. In fact, none of the statements of this section rely directly on Postulates (I) and (II). Instead, they use the following two assumptions, which are much weaker than Postulates (I) and (II). The first is the assumption that the sponge crossing probabilities at p_c are bounded away from one, that is,

$$1 - R_{n,3n}(p_c) > \varepsilon, \quad n \geq 1, \tag{4.1}$$

for some $\varepsilon > 0$, and the second is the assumption that (4.1) can be extended to $p > p_c$, provided $n \leq L_0(p)$. Actually, we only need the slightly weaker assumption that there are some constants $\varepsilon > 0$ and $\sigma_3 > 0$ such that

$$1 - R_{n,3n}(p) > \varepsilon \quad \text{for all} \quad p > p_c \quad \text{and all} \quad n \leq \sigma_3 L_0(p)\,. \tag{4.2}$$

To see that (4.1) follows from Postulates (I) and (II), we note that these postulates imply that $A_0(p) \to \infty$ as $p \downarrow p_c$, which in turn implies the statement (4.1). The bound (4.2) follows directly from Postulate (II), since, by the definition of $A_0(p)$,

$$1 - R_{r,3r}(p) > \varepsilon \quad \text{for} \quad r^{d-1} < A_0(p) \quad \text{and} \quad p > p_c.$$

By the equivalence of $A_0(p)$ and $L_0(p)^{d-1}$ (see Postulate (II)) this means that there exists some $\sigma_3 > 0$ such that (4.2) holds for $p > p_c$ and all $n \leq \sigma_3 L_0(p)$. We caution the reader that above p_c, the definition of the correlation length $L_0(p)$ in [BCKS99] is slightly different from the definition here (compare (2.17) in [BCKS99] to our equation (2.24)). However, as noted in Remark (vi) in [BCKS99], all results there remain valid for any definition of $L_0(p)$ above p_c that obeys Postulates (3.15) and (3.16) in [BCKS99]. While Postulate (3.16) of [BCKS99] is identical to our Postulate (III), Postulate (3.15) in [BCKS99] is slightly stronger than our assumption (4.2) — the former corresponds to (4.2) with $\sigma_3 = 1$. Here, we need only one result which uses Postulate (3.15), namely Theorem 3.6 of [BCKS99], which we cite to establish the last statement in our Proposition 4.8 below. However, a careful reading of the proof of Theorem 3.6 in equations (5.32) – (5.35) of [BCKS99] shows that actually only our weaker assumption (4.2) is needed.

4.1 General Moment Estimates.

The first lemma is a direct consequence of Postulate (IV). It is identical to Lemma 4.4 in [BCKS99].

Lemma 4.1. *If Postulate (IV) holds, then for $\beta > 1/\rho_1 - 1$ (and a fortiori for $\beta > d/2 - 1 = (d-2)/2$) there exists constants $C_1 = C(\beta, d)$ and $C_2 = C_2(d)$ such that*

$$\sum_{m=0}^{L} (m+1)^{\beta} \pi_m(p_c) \leq C_1 L^{\beta+1} \pi_L(p_c) \quad \text{if} \quad L \geq 1, \tag{4.3}$$

and

$$\sum_{m=0}^{L} (m+1)^{d-1} \pi_m^2(p_c) \leq C_2 L^d \pi_L^2(p_c) \quad \text{if} \quad L \geq 1. \tag{4.4}$$

The next lemma, which is identical to Lemma 6.1 in [BCKS99], gives a basic moment estimate. For $d = 2$ such an estimate was already given in [Ngu88].

Lemma 4.2. *Assume Postulate (IV) holds. Define*

$$V(L) := \text{number of sites in } \Lambda_L \text{ connected to } \partial\Lambda_{2L}. \tag{4.5}$$

Then for some constants C_i, it holds that

$$E_p V^k(L) \leq C_1 k! \left(C_2 L^d \pi_L(p_c) \right)^k \tag{4.6}$$

provided $p \leq p_c$, $k \geq 1$ and $L \geq 1$. Consequently

$$E_p \exp(tV(L)) \leq C_1 [1 - tC_2 L^d \pi_L(p_c)]^{-1} \tag{4.7}$$

whenever $p \leq p_c$ and $0 \leq t < [C_2 L^d \pi_L(p_c)]^{-1}$. When Postulates (III) and (IV) hold, then (4.6) and (4.7) remain valid for $p > p_c$ and $L \leq L_0(p)$.

The next proposition, which is one of the main technical results of [BCKS99] (Proposition 6.3 in [BCKS99]), follows from the above moment estimate Lemma 4.2. It is crucial for our proofs in Section 5.1 and 5.3.

Proposition 4.3.
i) Assume that Postulate (IV) holds. Then there exist constants C_i such that

$$Pr_p \left\{ W_{\Lambda_n}^{(1)} \geq xs(L_0(p)) \right\} \leq C_1 \left(\frac{n}{L_0(p)} \right)^d e^{-C_2 x} \tag{4.8}$$

if $x \geq 0$*,* $n \geq L_0(p)$*, and* $p < p_c$*. In particular*

$$Pr_p\left\{W^{(1)}_{\Lambda_n} \geq ys(L_0(p_n))\log\left(\frac{n}{L_0(p_n)}\right)\right\} \to 0 \tag{4.9}$$

if $y > d/C_2$ *and* $g(p_n, p) \to -\infty$*.*

ii) Assume that Postulate (IV) holds, and if $p > p_c$*, that also Postulate (III) holds. Then there exist constants* C_i *such that*

$$Pr_p\left\{W^{(1)}_{\Lambda_n} \geq xs(n)\right\} \leq C_1 e^{-C_2 x} \quad \text{if} \quad x \geq 0 \text{ and } n \leq L_0(p). \tag{4.10}$$

iii) Assume that Postulates (III) and (IV) hold. Then there exist constants C_i *such that*

$$Pr_p\left\{W^{(1)}_{\Lambda_n} \geq xs(L_0(p))\right\} \leq C_1\left(\frac{n}{L_0(p)}\right)^d \exp\left[-C_2 x + C_3\left(\frac{n}{L_0(p)}\right)^d\right] \tag{4.11}$$

if $x \geq 0$*,* $n \geq L_0(p)$ *and* $p > p_c$*.*

The next lemma summarize several additional results which follow from Postulate (IV). To state it, we introduce the diameter of a cluster $\mathcal{C}$ as

$$\operatorname{diam}(\mathcal{C}) = \max_{v,w\in\mathcal{C}} |v-w|_\infty. \tag{4.12}$$

Lemma 4.4. *Assume that Postulate (IV) holds. Then there exist constants* C_i *such that*

$$Pr_p\{\operatorname{diam}(\mathcal{C}(\mathbf{0})) \geq xL_0(p)\} \leq C_1\pi_{L_0(p)}(p)e^{-C_2 x} \quad \text{if} \quad x \geq 2 \text{ and } p < p_c, \tag{4.13}$$

and

$$Pr_p\{\exists \text{ cluster } \mathcal{C} \text{ in } \Lambda_n \text{ with } \operatorname{diam}(\mathcal{C}) \leq yn \text{ and } |\mathcal{C}| \geq xs(n)\} \leq C_1 y^{-d} e^{-C_2 x/y^{d/2}} \tag{4.14}$$

if $x \geq 0$*,* $0 < y \leq 1$*,* $p \leq p_c$ *and* $4/y \leq n \leq L_0(p)$*.*

Proof. The bound (4.14) was proved in [BCKS99], see Remark (xiii) at the end of Section 6 in [BCKS99].

To prove (4.13) we note that for $x \geq 2$,

$$\begin{aligned}
Pr_p\{\operatorname{diam}(\mathcal{C}(\mathbf{0})) \geq xL_0(p)\} &\leq Pr_p\{\mathbf{0} \leftrightarrow \partial B_{L_0(p)} \text{ and } \partial B_{L_0(p)} \text{ is connected to at least} \\
&\qquad \lfloor x/2 \rfloor \text{ distinct boxes } B_{L_0(p)}(\mathbf{j}L_0(p)),\ \mathbf{j} \in 2\mathbb{Z}^d \setminus \{\mathbf{0}\}\} \\
&= \pi_{L_0(p)}(p) Pr_p\{\partial B_{L_0(p)} \text{ is connected to at least} \\
&\qquad \lfloor x/2 \rfloor \text{ distinct boxes } B_{L_0(p)}(\mathbf{j}L_0(p)),\ \mathbf{j} \in 2\mathbb{Z}^d \setminus \{\mathbf{0}\}\}
\end{aligned}$$

(see (2.11) for the definition of $B_n(v)$). As in the proof of Proposition 6.3 (ii) of [BCKS99], (more precisely, as in the proof of the bound (6.39) in [BCKS99]), the renormalized Peierls argument of Theorem 5.1 in [Kes82] shows that for suitable constants C_1, C_2 the probability

$$\begin{aligned}
&Pr_p\{\partial B_{L_0(p)} \text{ is connected to at least} \\
&\qquad \lfloor x/2 \rfloor \text{ distinct boxes } B_{L_0(p)}(\mathbf{j}L_0(p)),\ \mathbf{j} \in 2\mathbb{Z}^d \setminus \{\mathbf{0}\}\}
\end{aligned}$$

is bounded above by $C_1 e^{-C_2 x}$. $\square$

4.2 Some Important Scaling Properties.

In this subsection we state a number of properties of the functions π_n, P_s and $\chi(p)$, most of which have already been proved in [BCKS99]. The first lemma provides an upper bound for $\pi_m(p_c)/\pi_n(p_c)$ which complements the lower bound of Postulate (IV).
5.1

Lemma 4.5. *i) There are constants* $C_1 < \infty$ *and* $C_2 > 0$ *such that*

$$\frac{\pi_n(p)}{\pi_{L_0(p)}(p)} \leq C_1 e^{-C_2 n/L_0(p)} \quad \text{if} \quad p < p_c \quad \text{and} \quad n \geq L_0(p)\,. \tag{4.15}$$

ii) Assume that (4.1) holds for some $\varepsilon > 0$*. Then*

$$Pr_{p_c}\{\partial B_n(\mathbf{0}) \leftrightarrow \partial B_{3n}(\mathbf{0})\} \leq 1 - \varepsilon^{2d} \quad \text{if} \quad n \geq 1. \tag{4.16}$$

iii) Assume that (4.1) holds for some $\varepsilon > 0$*. Then there exist constants* $C_1, \rho_2 < \infty$ *such that*

$$\frac{\pi_m(p_c)}{\pi_n(p_c)} \leq C_1 \Big(\frac{m}{n}\Big)^{-1/\rho_2} \quad \text{if} \quad m \geq n \geq 1. \tag{4.17}$$

Proof. Statements i) and iii) are the content of Theorem 3.8 of [BCKS99]. To prove ii), we show that for any $p \in [0,1]$ and any $n \geq 1$, one has

$$Pr_p\{\partial B_n(\mathbf{0}) \not\leftrightarrow \partial B_{3n}(\mathbf{0})\} \geq [1 - R_{2n,6n}(p)]^{2d}. \tag{4.18}$$

Indeed, by the definition of $R_{n,m}$, the probability that there is no occupied crossing in the 1-direction of the block

$$[n, 3n] \times [-3n, 3n]^{d-1} \tag{4.19}$$

is equal to $1 - R_{2n,6n}$. The cube $B_{3n}(\mathbf{0})$ is the union of $B_n(\mathbf{0})$ and the block in (4.19) plus $2d-1$ more blocks congruent to the block in (4.19). Let F_n be the event that none of these $2d$ blocks congruent to (4.19) has an occupied crossing in the short direction. Obviously, the event F_n implies that $\partial B_n(\mathbf{0})$ is not connected to $\partial B_{3n}(\mathbf{0})$, so that the probability on the left hand side of (4.18) is bounded from below by the probability of F_n. Since $Pr_p\{F_n\}$ is at least $[1 - R_{2n,6n}(p)]^{2d}$ by the Harris–FKG inequality, the bound (4.18) follows. □

The next proposition summarizes the results of Theorem 3.7 and the first statement of Theorem 3.4 in [BCKS99]. Assuming existence of the critical exponents ρ and δ, the first statement implies the hyperscaling relation $d\rho = \delta + 1$. The second statement is the analogue of Postulate (III) for $P_{\geq s}(p)$.

Proposition 4.6. *Assume that (4.1) holds for some* $\varepsilon > 0$ *and that Postulate (IV) holds. Then there exists constants* $C_1 > 0$ *and* $C_2 < \infty$ *such that*

$$C_1 \pi_n(p_c) \leq P_{\geq s(n)}(p_c) \leq C_2 \pi_n(p_c)\,. \tag{4.20}$$

If Postulate (III) holds as well, then there exists constants $C_3 > 0$, $C_4 < \infty$ *and* $0 < \sigma_0 = \sigma_0(\varepsilon, d) \leq 1$ *such that*

$$C_3 P_{\geq s(n)}(p_c) \leq P_{\geq s(n)}(p) \leq C_4 P_{\geq s(n)}(p_c) \quad \text{if} \quad n \leq \sigma_0 L_0(p)\,. \tag{4.21}$$

Our last two propositions in this section summarizes the results of several theorems in [BCKS99], namely Theorem 3.5, Theorem 3.6 and Theorem 3.9. Proposition 4.8 in particular has two upper bounds complementing lower bounds in the postulates, and a hyperscaling relation. Assuming the existence of the corresponding exponents, this relation implies $\gamma = (d - 2/\rho)\nu$.

Lemma 4.7. *Assume Postulate (IV) holds. Then there exist constants* $0 < C_i < \infty$ *such that*

$$\frac{P_{\geq xs(L_0(p))}(p)}{\pi_{L_0(p)}(p_c)} \leq C_1 e^{-C_2 x} \quad \text{if} \quad p < p_c \text{ and } x \geq 1. \tag{4.22}$$

Proposition 4.8. *Assume that (4.1) is valid for some* $\varepsilon > 0$*, and that Postulates (III) and (IV) hold. Then there exist constants* $0 < C_i < \infty$ *such that, with* σ_0 *as in Proposition 4.6, it holds that*

$$\frac{P_{\geq xs(L_0(p))}(p)}{P_{\geq s(\sigma_0 L_0(p))}(p)} \leq C_1 \exp[-C_2 x] \quad \text{if} \quad x \geq 1 \text{ and } p < p_c, \tag{4.23}$$

and

$$C_3 L_0(p)^d [\pi_{L_0(p)}(p_c)]^2 \leq \chi(p) \leq C_4 L_0(p)^d [\pi_{L_0(p)}(p_c)]^2, \quad p < p_c. \tag{4.24}$$

If (4.1) and (4.2) are valid for some $\varepsilon > 0$ *and some* $\sigma_3 > 0$*, and if Postulate (IV) holds, then there exists a constant* $C_5 > 0$ *such that*

$$C_5 L_0(p)^d [\pi_{L_0(p)}(p_c)]^2 \leq \chi^{\text{fin}}(p), \quad p > p_c. \tag{4.25}$$

5. Proof of the Theorems, Given the Postulates

In this section, we prove our principal results, Theorems 3.1–3.5. The section is divided into three subsections. These correspond to the proof of results within, above and below the scaling window: Theorem 3.1 i), Theorem 3.3 and Theorem 3.4 in Section 5.1, Theorem 3.1 iii) and Theorem 3.2 in Section 5.2, and finally, Theorem 3.1 ii) and Theorem 3.5 in Section 5.3.

5.1 Inside the Scaling Window.

We start this subsection with several lemmas and propositions concerning the numbers $N_\Lambda(s_1, s_2)$ and $\widetilde{N}_\Lambda(s_1, s_2)$ of clusters with size between s_1 and s_2, defined in (2.1) and (2.2). Although some of these results are very similar to the theorems we are finally going to prove, we give them as separate propositions, since this allows us to better keep track of which postulates are needed in which step.

At many points in this and the following subsections, we use the fact that, for an arbitrary configuration ω, and number α, it holds that

$$\sum_{i\geq 1}\big[W_\Lambda^{(i)}\big]^\alpha=\sum_{v\in\Lambda}\sum_{s\geq 1}s^{\alpha-1}I[|\mathcal{C}_\Lambda(v)|=s]. \tag{5.1}$$

This is obvious from the fact that in the right hand side, the sum of $I[|\mathcal{C}_\Lambda(w)|=s]$ over all point w in $\mathcal{C}_\Lambda(v)$ equals $sI[|\mathcal{C}_\Lambda(v)|=s]$. Taking expectations of (5.1) gives

$$\sum_{i\geq 1}E_p\big\{\big[W_\Lambda^{(i)}\big]^\alpha\big\}=\sum_{v\in\Lambda}\sum_{s\geq 1}s^{\alpha-1}Pr_p\left\{|\mathcal{C}_\Lambda(v)|=s\right\}. \tag{5.2}$$

This argument for $\alpha=1$ will be used in the proof of Proposition 5.5, but even more often will we use the special case $\alpha=0$, which says that the number of clusters of size s can be rewritten as

$$\big|\{i\mid W_\Lambda^{(i)}=s\}\big|=\sum_{v\in\Lambda}\frac{1}{s}I[|\mathcal{C}_\Lambda(v)|=s]. \tag{5.3}$$

These formulae and some variants form a basic relationship which allows us to relate estimates on the distributions of $W_\Lambda^{(i)}$ and $|\mathcal{C}(\mathbf{0})|$. We use the following consequence of (5.3):

$$E_p\big\{N_\Lambda(s_1,s_2)\big\}=\sum_{s=s_1}^{s_2}\sum_{v\in\Lambda}\frac{1}{s}Pr_p\big\{|\mathcal{C}_\Lambda(v)|=s\big\}. \tag{5.4}$$

In a similar way, we have

$$E_p\big\{\widetilde{N}_\Lambda(s_1,s_2)\big\}=\sum_{s=s_1}^{s_2}\sum_{v\in\Lambda}\frac{1}{s}Pr_p\big\{|\mathcal{C}_\Lambda(v)|=s,\,v\not\leftrightarrow\partial\Lambda\big\}. \tag{5.5}$$

We also need the corresponding representation for the expectation of $\widetilde{N}_\Lambda^2(s_1,s_2)$:

$$E_p\big\{\widetilde{N}_\Lambda^2(s_1,s_2)\big\}=\sum_{\substack{s_1\leq s\leq s_2\\ s_1\leq\tilde{s}\leq s_2}}\sum_{v,w\in\Lambda}\frac{1}{s\tilde{s}}Pr_p\big\{|\mathcal{C}_\Lambda(v)|=s,v\not\leftrightarrow\partial\Lambda,|\mathcal{C}_\Lambda(w)|=\tilde{s},w\not\leftrightarrow\partial\Lambda\big\}. \tag{5.6}$$

The next two lemmas will be useful in proving lower bounds for $W_\Lambda^{(i)}$.

Proposition 5.1. *Assume that (4.1) holds for some $\varepsilon>0$ and that Postulates (III) and (IV) hold. Then there exist constants $0<C_i<\infty$ and $1\leq\sigma_1<\infty$ such that*

$$C_1\Big(\frac{n}{m}\Big)^d\leq E_p\big\{\widetilde{N}_{\Lambda_n}(s(m),s(km))\big\}\leq E_p\big\{N_{\Lambda_n}(s(m),s(km))\big\}\leq C_2\Big(\frac{n}{m}\Big)^d \tag{5.7}$$

provided $\sigma_1 m\leq\min\{L_0(p),n\}$ and $k\geq\sigma_1$.

Proof. For brevity we write Λ instead of Λ_n. We start with the upper bound. Using the representation (5.4) and bounding the factor $1/s$ in (5.4) by $1/s(m)$, we get

$$\begin{aligned} E_p\{N_\Lambda(s(m), s(km))\} &\leq \frac{1}{s(m)} \sum_{v\in\Lambda} \sum_{s\geq s(m)} Pr_p\{|\mathcal{C}_\Lambda(v)| = s\} \\ &= \frac{1}{s(m)} \sum_{v\in\Lambda} Pr_p\{|\mathcal{C}_\Lambda(v)| \geq s(m)\} \\ &\leq \frac{(2n)^d}{s(m)} P_{\geq s(m)}(p)\,, \end{aligned} \tag{5.8}$$

where in the last step we used the definition (2.4) of $P_{\geq s(m)}(p)$ and the fact that $|\mathcal{C}_\Lambda(v)| \leq |\mathcal{C}(v)|$. Without loss of generality we shall take $\sigma_1 \geq 1/\sigma_0 \geq 1$, where σ_0 is the constant of Proposition 4.6. Then $\sigma_1 m \leq L_0(p)$ implies $m \leq \sigma_0 L_0(p)$, and we may use Proposition 4.6 to bound the right hand side of (5.8). We get for some finite constant C_2

$$\frac{(2n)^d}{s(m)} P_{\geq s(m)}(p) \leq C_2 \frac{(2n)^d}{s(m)} \pi_m(p_c) = C_2 \left(\frac{n}{m}\right)^d. \tag{5.9}$$

The estimates (5.8) and (5.9) imply the upper bound.

To prove the lower bound, we use that Postulate (IV) implies that

$$\frac{s(\ell)}{s(\ell')} \geq D_3 \left(\frac{\ell}{\ell'}\right)^{d/2} \quad \text{if} \quad \ell \geq \ell' \geq 1, \tag{5.10}$$

so that in particular $s(\ell) \geq s(\ell')$ whenever $\ell/\ell' \geq D_3^{-2/d}$. We conclude that for any choice of $\widetilde{k} \geq 1$ we can find a $\sigma_1 \geq \widetilde{k}(1 + 1/\sigma_0)$ such that $s(km) \geq s(\widetilde{k}m)$ for all $k \geq \sigma_1$. It then follows from (5.5) that for $k \geq \sigma_1$

$$\begin{aligned} E_p\{\widetilde{N}_\Lambda(s(m), s(km))\} &\geq E_p\{\widetilde{N}_\Lambda(s(m), s(\widetilde{k}m) - 1)\} \\ &\geq \sum_{s=s(m)}^{s(\widetilde{k}m)-1} \sum_{v\in\Lambda_{\frac{n}{2}}} \frac{1}{s} Pr_p\{|\mathcal{C}_\Lambda(v)| = s, x \not\leftrightarrow \partial\Lambda\} \\ &= \sum_{s=s(m)}^{s(\widetilde{k}m)-1} \sum_{v\in\Lambda_{\frac{n}{2}}} \frac{1}{s} Pr_p\{|\mathcal{C}(v)| = s, v \not\leftrightarrow \partial\Lambda\}\,, \end{aligned} \tag{5.11}$$

where in the second step we bounded the sum over $\Lambda = \Lambda_n$ from below by a sum over $\Lambda_{\frac{n}{2}}$.

Bounding the factor $1/s$ in (5.11) from below by $1/s(\widetilde{k}m)$, we get

$$\begin{aligned} E_p\Big\{\widetilde{N}_\Lambda(s(m),s(km))\Big\} \\ &\geq \frac{1}{s(\widetilde{k}m)}\sum_{v\in\Lambda_{\frac{n}{2}}} Pr_p\left\{s(m)\leq|\mathcal{C}(v)|<s(\widetilde{k}m), v\not\leftrightarrow\partial\Lambda\right\} \\ &= \frac{1}{s(\widetilde{k}m)}\sum_{v\in\Lambda_{\frac{n}{2}}}\Big[Pr_p\left\{s(m)\leq|\mathcal{C}(v)|<s(\widetilde{k}m)\right\} \\ &\qquad - Pr_p\left\{s(m)\leq|\mathcal{C}(v)|<s(\widetilde{k}m), v\leftrightarrow\partial\Lambda\right\}\Big] \\ &\geq \frac{1}{s(\widetilde{k}m)}\sum_{v\in\Lambda_{\frac{n}{2}}}\Big[Pr_p\left\{s(m)\leq|\mathcal{C}(v)|<s(\widetilde{k}m)\right\} - \pi_{n/2}(p)\Big] \\ &\geq \frac{(n-2)^d}{s(\widetilde{k}m)}\Big[P_{\geq s(m)}(p) - P_{\geq s(\widetilde{k}m)}(p) - \pi_{n/2}(p)\Big]. \end{aligned} \tag{5.12}$$

Since $n\geq\sigma_1 m\geq\widetilde{k}m$ by the assumption $\sigma_1 m\leq\min\{L_0(p),n\}$, we obtain

$$E_p\left\{\widetilde{N}_\Lambda(s(m),s(km))\right\} \geq \frac{(n-2)^d}{s(\widetilde{k}m)}\Big[P_{\geq s(m)}(p) - P_{\geq s(\widetilde{k}m)}(p) - \pi_{\widetilde{k}m/2}(p)\Big]. \tag{5.13}$$

Again by the assumption $\sigma_1 m\leq\min\{L_0(p),n\}$, we have $m\leq\widetilde{k}m\leq(\widetilde{k}/\sigma_1)L_0(p)\leq\sigma_0 L_0(p)$. We therefore may use Proposition 4.6 in conjunction with Postulate (III) and the bound $\pi_{\widetilde{k}m}(p_c)\leq\pi_{\widetilde{k}m/2}(p_c)$ to conclude that

$$E_p\left\{\widetilde{N}_\Lambda(s(m),s(km))\right\} \geq \frac{(n-2)^d}{s(\widetilde{k}m)}\Big[C_3\pi_m(p_c) - C_4\pi_{\widetilde{k}m/2}(p_c)\Big], \tag{5.14}$$

for suitable constants $C_3, C_4\in(0,\infty)$ which depend only on the constants in Proposition 4.6, but not on the choice of $\widetilde{k}$. Finally we appeal to Lemma 4.5 iii) to fix $\widetilde{k}$ so large that $C_4\pi_{\widetilde{k}m/2}(p_c)\leq\frac{1}{2}C_3\pi_{m(p_c)}$. Here $\widetilde{k}$ depends only on C_4/C_3 and the constants in Lemma 4.5 iii); also $\widetilde{k}$ determines the value to take for σ_1. We then get

$$E_p\left\{\widetilde{N}_\Lambda(s(m),s(km))\right\} \geq \frac{(n-2)^d}{s(\widetilde{k}m)}\tfrac{1}{2}C_3\pi_m(p_c). \tag{5.15}$$

From $s(\widetilde{k}m)\leq\widetilde{k}^d s(m)$ we then conclude that for $n\geq 4$

$$E_p\left\{\widetilde{N}_\Lambda(s(m),s(km))\right\} \geq C_1\frac{(2n)^d}{s(m)}\pi_m(p_c) = C_1\left(\frac{n}{m}\right)^d, \tag{5.16}$$

where $C_1 = 2^{-2d-1}\widetilde{k}^{-d}C_3$. This proves the lower bound when $n\geq 4$. If we choose σ_1 large enough, then $1\leq n<4$ is ruled out by $\sigma_1\leq\sigma_1 m\leq n$. □

Proposition 5.2. *Assume that (4.1) holds for some $\varepsilon > 0$ and that Postulates (III) and (IV) hold. Then there is a constant $C_3 < \infty$, such that*

$$\frac{Var\{\widetilde{N}_{\Lambda_n}(s(m), s(km))\}}{E_p\{\widetilde{N}_{\Lambda}(s(m), s(km))\}^2} \leq C_3\,, \tag{5.17}$$

provided $\sigma_1 m \leq \min\{L_0(p), n\}$, $k \geq \sigma_1$. Here σ_1 is the constant of Proposition 5.1.

Proof. Again we write Λ for Λ_n. We first will prove that for arbitrary $s_1, s_2 \in \mathbb{N}$, $s_1 \leq s_2$, and $p \in (0,1)$,

$$E_p\left\{[\widetilde{N}_{\Lambda}(s_1, s_2)]^2\right\} \;\leq\; E_p\left\{\widetilde{N}_{\Lambda}(s_1, s_2)\right\}\left[1 + \frac{(2n)^d}{s_1} P_{\geq s_1}(p)\right]. \tag{5.18}$$

We need some notation. We denote the set of bonds with both endpoints in Λ by $B(\Lambda)$, and the set of bonds with both endpoints in $\Lambda \setminus \partial\Lambda$ by $\widetilde{B}(\Lambda)$. Let B be a subset of $B(\Lambda)$. With a slight abuse of notation, we say that v is a point in B if v is an endpoint of one of the bonds in B. We write B is occupied (vacant) for the event that *all* bonds in $B \subset B(\Lambda)$ are occupied (respectively, vacant). Given $v \in \Lambda$, we denote the set of all connected subsets $B \subset \widetilde{B}(\Lambda)$ that contain the point v by $\mathcal{B}_v(\Lambda)$. Again with a slight abuse of notation, we denote the number of points in a cluster $B \subset \mathcal{B}_v(\Lambda)$ by $|B|$. Finally, we write $\partial^\star_\Lambda B$ for the set of all bonds $b \in B(\Lambda) \setminus B$ which share an endpoint with a bond $b' \in B$.

Using equation (5.6), we rewrite the left-hand side of (5.18) as

$$\begin{aligned} &E_p\{\widetilde{N}_{\Lambda}(s_1, s_2)^2\} \\ &\quad= \sum_{v,w\in\Lambda} \sum_{\substack{B\in\mathcal{B}_v(\Lambda)\\ s_1\leq|B|\leq s_2}} \sum_{\substack{\widetilde{B}\in\mathcal{B}_w(\Lambda)\\ s_1\leq|\widetilde{B}|\leq s_2}} \frac{Pr_p\{B\cup\widetilde{B} \text{ is occupied}, \partial^\star_\Lambda B \cup \partial^\star_\Lambda \widetilde{B} \text{ is vacant}\}}{|B|\,|\widetilde{B}|}\,. \end{aligned} \tag{5.19}$$

Next we observe that the event on the right-hand side cannot occur if $B \leftrightarrow \widetilde{B}$ and $B \neq \widetilde{B}$ in Λ, because in this case some occupied bond in $B \cup \widetilde{B} \cup$ (a suitable path from B to $\widetilde{B}$) also lies in $\partial^\star_\Lambda B \cup \partial^\star_\Lambda \widetilde{B}$. As a consequence, the right-hand side decomposes into two terms: the term

$$\begin{aligned} &\sum_{v,w\in\Lambda} \sum_{\substack{B\in\mathcal{B}_v(\Lambda)\cap\mathcal{B}_w(\Lambda)\\ s_1\leq|B|\leq s_2}} \frac{Pr_p\{B \text{ is occupied}, \partial^\star_\Lambda B \text{ is vacant}\}}{|B|^2} \\ &\quad= \sum_{v\in\Lambda} \sum_{\substack{B\in\mathcal{B}_v(\Lambda)\\ s_1\leq|B|\leq s_2}} \frac{Pr_p\{B \text{ is occupied}, \partial^\star_\Lambda B \text{ is vacant}\}}{|B|} \\ &\quad= E_p\left\{\widetilde{N}_{\Lambda}(s_1, s_2)\right\} \end{aligned} \tag{5.20}$$

and the term

$$\sum_{v,w\in\Lambda}\sum_{\substack{B\in\mathcal{B}_v(\Lambda)\\ s_1\le|B|\le s_2}}\sum_{\substack{\widetilde{B}\in\mathcal{B}_w(\Lambda)\\ s_1\le|\widetilde{B}|\le s_2,\widetilde{B}\not\leftrightarrow B}}\frac{Pr_p\{B\cup\widetilde{B}\text{ is occupied},\partial^\star_\Lambda B\cup\partial^\star_\Lambda\widetilde{B}\text{ is vacant}\}}{|B|\,|\widetilde{B}|} \tag{5.21}$$

By using the second decoupling inequality of [BC96], or, alternatively, the van den Berg-Kesten inequality [BK85] we see that the last sum equals

$$\begin{aligned}
&\sum_{v,w\in\Lambda}\sum_{\substack{B\in\mathcal{B}_v(\Lambda)\setminus\mathcal{B}_w(\Lambda)\\ s_1\le|B|\le s_2}}\sum_{\substack{\widetilde{B}\in\mathcal{B}_w(\Lambda)\\ s_1\le|\widetilde{B}|\le s_2}}\frac{Pr_p\{B\cup\widetilde{B}\text{ is occupied},\partial^\star_\Lambda B\cup\partial^\star_\Lambda\widetilde{B}\text{ is vacant}\}}{|B|\,|\widetilde{B}|}\\
&\le\sum_{v,w\in\Lambda}\frac{1}{s_1}\sum_{\substack{B\in\mathcal{B}_v(\Lambda)\setminus\mathcal{B}_w(\Lambda)\\ s_1\le|B|\le s_2}}\frac{Pr_p\{B\text{ is occupied},\partial^\star_\Lambda B\text{ is vacant},|\mathcal{C}_\Lambda(w)|\ge s_1\}}{|B|}\\
&\le\sum_{v,w\in\Lambda}\frac{1}{s_1}\sum_{\substack{B\in\mathcal{B}_v(\Lambda)\setminus\mathcal{B}_w(\Lambda)\\ s_1\le|B|\le s_2}}\frac{Pr_p\{B\text{ is occupied},\partial^\star_\Lambda B\text{ is vacant}\}\ \ Pr_p\{|\mathcal{C}_\Lambda(w)|\ge s_1\}}{|B|}\\
&\le E_p\left\{\widetilde{N}_\Lambda(s_1,s_2)\right\}\sum_{w\in\Lambda}\frac{Pr_p\{|\mathcal{C}_\Lambda(w)|\ge s_1\}}{s_1}\,.
\end{aligned} \tag{5.22}$$

Combining the two terms (5.20) and (5.22), and observing that $Pr_p\{|\mathcal{C}_\Lambda(w)|\ge s_1\}\le Pr_p\{|\mathcal{C}(w)|\ge s_1\}$, we obtain (5.18). The bound (5.17) now follows from (5.18), (5.9) and the lower bound in (5.7). □

The next proposition is a consequence of Proposition 5,1, Proposition 5.2 and the fact that

$$\widetilde{N}_\Lambda(s(m),s(km))\ge\widetilde{N}_{\Lambda_1}(s(m),s(km))+\widetilde{N}_{\Lambda_2}(s(m),s(km)),$$

provided $\Lambda_1\subset\Lambda$ and $\Lambda_2=\Lambda\setminus\Lambda_1$.

Proposition 5.3. *Assume that (4.1) holds for some $\varepsilon>0$ and that Postulates (III) and (IV) hold. Then there are constants $C_4,C_5>0$ such that*

$$Pr_p\left\{\widetilde{N}_{\Lambda_n}(s(m),s(km))\ge C_4\left(\frac{n}{m}\right)^d\right\}\ge 1-C_5\left(\frac{m}{n}\right)^d \tag{5.23}$$

provided $\sigma_1 m\le\min\{L_0(p),n\}$ and $k\ge\sigma_1$. Here σ_1 is the constant of Proposition 5.1.

Proof. Let $\widetilde{k}=\lfloor n/\lceil\sigma_1 m\rceil\rfloor$ be the largest integer less than or equal to $n/\lceil\sigma_1 m\rceil$, and $\widetilde{n}=\widetilde{k}\lceil\sigma_1 m\rceil$. Note that then $\sigma_1 m\le\widetilde{n}\le n$. Since $\widetilde{N}_\Lambda(s(m),s(km))$ is increasing in Λ, i.e.,

$$\widetilde{N}_\Lambda(s(m),s(km))\ge\widetilde{N}_{\widetilde{\Lambda}}(s(m),s(km))\text{ if }\widetilde{\Lambda}\subset\Lambda\,, \tag{5.24}$$

we get that for $\Lambda = \Lambda_n, \widetilde{\Lambda} = \Lambda_{\widetilde{n}}$,

$$Pr_p\Big\{\widetilde{N}_\Lambda(s(m), s(km)) \geq C_4\Big(\frac{n}{m}\Big)^d\Big\} \geq Pr_p\Big\{\widetilde{N}_{\widetilde{\Lambda}}(s(m), s(km)) \geq C_4\Big(\frac{n}{m}\Big)^d\Big\}. \tag{5.25}$$

Next we note that $\widetilde{\Lambda}$ contains $\widetilde{k}^d$ disjoint subvolumes $\Lambda^{(i)}$ of size $(2\lceil\sigma_1 m\rceil)^d$, and introduce the random variable

$$X = \sum_{i=1}^{\widetilde{k}^d} \widetilde{N}_{\Lambda^{(i)}}(s(m), s(km))\,. \tag{5.26}$$

Since $X \leq \widetilde{N}_{\widetilde{\Lambda}}(s(m), s(km)) \leq \widetilde{N}_\Lambda(s(m), s(km))$, we have

$$Pr_p\Big\{\widetilde{N}_\Lambda(s(m), s(km)) \geq C_4\Big(\frac{n}{m}\Big)^d\Big\} \geq Pr_p\Big\{X \geq C_4\Big(\frac{n}{m}\Big)^d\Big\}. \tag{5.27}$$

Observing that the random variables $\widetilde{N}_{\Lambda^{(i)}}(s(m), s(km))$ in (5.26) are i.i.d. and using Proposition 5.2, we have

$$\frac{\mathrm{Var}X}{(E_pX)^2} = \frac{1}{\widetilde{k}^d}\frac{\mathrm{Var}\{\widetilde{N}_{\Lambda^{(1)}}(s(m), s(km))\}}{E_p\{\widetilde{N}_{\Lambda^{(1)}}(s(m), s(km))\}^2} \leq \frac{C_6}{\widetilde{k}^d}\,. \tag{5.28}$$

Noting that

$$Pr_p\big\{X \leq \tfrac{1}{2}E_pX\big\} \leq Pr_p\big\{|X - E_pX|^2 \geq \tfrac{1}{4}(E_pX)^2\big\} \leq \frac{4\mathrm{Var}X}{(E_pX)^2}\,, \tag{5.29}$$

we find that

$$Pr_p\left\{X \geq \tfrac{1}{2}E_pX\right\} \geq 1 - \frac{4C_6}{\widetilde{k}^d} \geq 1 - C_5\Big(\frac{m}{n}\Big)^d\,, \tag{5.30}$$

where $C_5 = (4\sigma_1)^d 4C_6$ (note that $1/\widetilde{k} = \lfloor n/\lceil\sigma_1 m\rceil\rfloor^{-1} \leq 4\sigma_1 m/n$). Using finally the lower bound

$$E_pX = \widetilde{k}^d E_p\left\{\widetilde{N}_{\Lambda^{(1)}}(s(m), s(km))\right\} \geq C_1(\widetilde{k}\sigma_1)^d = C_1\left(\frac{\widetilde{n}}{m}\right)^d,$$

which comes from (5.7), we obtain the desired bound (5.23), provided $C_4 > 0$ is chosen small enough. □

Proposition 5.4. *Suppose that Postulates (III) and (IV) hold, and that (4.1) and (4.2) are valid for some $\varepsilon > 0$ and some $\sigma_3 > 0$. Then there are strictly positive constants C_1 and σ_4 such that*

$$Pr_p\left\{W^{(1)}_{\Lambda_n} \leq s(m)\right\} \geq C_1^{1+(n/m)^d}\,, \tag{5.31}$$

provided $m \leq \sigma_4 L_0(p)$.

Proof. It follows from (4.10) and (5.10) that there exists a constant $\sigma_4 > 0$ such that

$$Pr_p\left\{W^{(1)}_{\Lambda_{3r}} \leq s(m)\right\} \geq \frac{1}{2} \qquad \text{if} \qquad r \leq \sigma_4 m \text{ and } r \leq \frac{1}{3}L_0(p). \tag{5.32}$$

In addition, it follows (4.18) that

$$Pr_p\{v \not\leftrightarrow \partial\Lambda_{3r} \text{ for all } v \in \Lambda_r\} \geq [1 - R_{r,3r}(p)]^{2d}. \tag{5.33}$$

For $p \leq p_c$,

$$1 - R_{r,3r}(p) \geq 1 - R_{r,3r}(p_c) > \varepsilon$$

(see (4.1)). We still have $1 - R_{r,3r}(p) > \varepsilon$ for $p > p_c$ and $r \leq \sigma_3 L_0(p)$, by virtue of (4.2). Consequently, as in (4.16),

$$Pr_p\{v \not\leftrightarrow \partial B_{3r} \text{ for all } v \in \Lambda_r\} \geq \varepsilon^{2d} \qquad \text{if} \qquad r \leq \sigma_3 L_0(p). \tag{5.34}$$

Using the Harris-FKG inequality we obtain from (5.32) and (5.34) that

$$\begin{aligned} &Pr_p\big\{|\mathcal{C}(v)| \leq s(m) \text{ for all } v \in \Lambda_r\big\} \\ &\geq Pr_p\Big\{W_{3r}^{(1)} \leq s(m) \text{ and } v \not\leftrightarrow \partial B_{3r} \text{ for all } v \in \Lambda_r\Big\} \\ &\geq \frac{1}{2}\varepsilon^{2d} \qquad \text{if} \qquad r \leq (\sigma_3 \wedge 1/3)L_0(p) \wedge \sigma_4 m. \end{aligned} \tag{5.35}$$

We are now ready to prove (5.31) for arbitrary n. We first estimate

$$Pr_p\Big\{W_\Lambda^{(1)} \leq s(m)\Big\} \geq Pr_p\Big\{|\mathcal{C}(v)| \leq s(m) \text{ for all } v \in \Lambda\Big\} \tag{5.36}$$

and note that the right-hand side of (5.36) is decreasing in Λ. Let $m \leq \sigma_4 L_0(p)$ and choose $0 < \sigma_5 \leq \sigma_4$ such that $\sigma_4\sigma_5 \leq (\sigma_3 \wedge 1/3)$. Then choose an integer $r \geq 1$ in $[\sigma_5 m/2, \sigma_5 m]$; if this is not possible, because $\sigma_5 m < 1$, then take $r = 1$. For this choice of r

$$Pr_p\{|\mathcal{C}(v)| \leq s(m) \text{ for all } v \in \Lambda_r\} \geq C_1 > 0$$

for some constant C_1, by virtue of (5.35). If $n < r$, then this already implies (5.31). Otherwise, choose an integer k such that $n \leq \widetilde{n} := kr \leq 2n$. We then get

$$Pr_p\Big\{W_{\Lambda_n}^{(1)} \leq s(m)\Big\} \geq Pr_p\Big\{\bigcap_{v \in \Lambda_{\widetilde{n}}} \{|\mathcal{C}(v)| \leq s(m)\}\Big\}. \tag{5.37}$$

Decomposing $\Lambda_{\widetilde{n}}$ into k^d subvolumes $\Lambda^{(i)}$ of diameter $2r$, and using the Harris-FKG inequality for the intersection of the events $\cap_{v\in\Lambda^{(i)}}\{|C(v)| \leq s(m)\}$, we obtain

$$Pr_p\Big\{W_{\Lambda_n}^{(1)} \leq s(m)\Big\} \geq \Big[Pr_p\{\bigcap_{v\in\Lambda_r} \{|\mathcal{C}(v)| \leq s(m)\}\}\Big]^{k^d} \geq C_1^{k^d}. \tag{5.38}$$

The proof is concluded by observing that $k \leq 2n/r \leq 4n/(\sigma_5 m)$. □

Proof of Theorem 3.1 i). For this proof we only use (4.1) and Postulates (III) and (IV). As before, abbreviate Λ_n to Λ. Since $\limsup_{n\to\infty} |g(p_n, n)| < \infty$, we have

$$n \le \lambda L_0(p_n) \text{ for all } n \ge n_1\,, \tag{5.39}$$

where λ and n_1 are finite constants depending on the sequence $\{p_n\}$.

The fact that $E_p\left\{W_\Lambda^{(1)}\right\}/s(n)$ is bounded above is immediate from Proposition 4.3. If $n \le L_0(p_n)$ then (4.10) suffices. If $L_0(p_n) \le n \le \lambda L_0(p_n)$, then we use (4.8) or (4.11) plus the fact that $s(n) \ge D_3 s(L_0(p_n))$ (by (5.10)). Note that this proof only requires Postulates (III) and (IV), and does not rely on the assumptions (4.1).

In order to complete the proof, we need lower bounds on $E_p\left\{W_\Lambda^{(i)}\right\}$. To this end, we first note that Proposition 5.3 implies that for any $\delta > 0$ there are constants $1 \le \sigma^{(i)} = \sigma^{(i)}(\lambda, \delta) < \infty$ such that

$$Pr_p\left\{W_{\Lambda_n}^{(i)} \ge s(m)\right\} \ge 1 - \delta \tag{5.40}$$

provided $\sigma^{(i)} m \le n \le \lambda L_0(p)$. Indeed, choose $\sigma^{(i)}(\lambda, \delta) \ge \sigma_1$ (with the constant σ_1 as in Proposition 5.1) so large that i) $\sigma^{(i)} m \le \lambda L_0(p)$ implies $\sigma_1 m \le L_0(p)$, ii) $C_4(\sigma^{(i)})^d \ge i$, and iii) $C_5(\sigma^{(i)})^{-d} \le \delta$, where C_4, C_5 are as in Proposition 5.3. Then for $\sigma^{(i)} m \le n \le \lambda L_0(p)$, we get

$$\begin{aligned} Pr_p\{W_\Lambda^{(i)} \ge s(m)\} = Pr_p\{N_\Lambda(s(m), \infty) \ge i\} &\ge Pr_p\{\widetilde{N}_\Lambda(s(m), s(\sigma_1 m)) \ge i\} \\ &\ge Pr_p\Big\{\widetilde{N}_\Lambda(s(m), s(\sigma_1 m)) \ge C_4\Big(\frac{n}{m}\Big)^d\Big\}\,, \end{aligned} \tag{5.41}$$

where we used that $\sigma^{(i)} m \le n$ implies $C_4(n/m)^d \ge i$ in the last step. Combined with Proposition 5.3 and the fact that the assumption $\sigma^{(i)} m \le n$ implies $C_5(m/n)^d \le \delta$ by our choice of $\sigma^{(i)}$, the bound (5.41) implies (5.40).

In order to prove a lower bound on $\liminf_{n\to\infty} E_{p_n}\{W_{\Lambda_n}^{(i)}\}$, we now assume that $n \ge n_1^{(i)} := \max\{n_1, \sigma^{(i)}\}$, where n_1 and λ are the constants from (5.39), and $\sigma^{(i)} = \sigma^{(i)}(\lambda, \frac{1}{2})$. Choosing $m = \lfloor n/\sigma^{(i)} \rfloor$, we have $m \ge 1$ and $\sigma^{(i)} m \le n \le \lambda L_0(p_n)$. Thus, by (5.40)

$$E_{p_n}\{W_{\Lambda_n}^{(i)}\} \ge \tfrac{1}{2} s(m)\,. \tag{5.42}$$

Since $m \le n/\sigma^{(i)} \le m + 1 \le 2m$ by the definition of m, we have

$$s(n)/s(m) \le (n/m)^d \le (2\sigma^{(i)})^d, \tag{5.43}$$

and hence $s(m) \ge s(n)(2\sigma^{(i)})^{-d}$. Thus, with $C_1^{(i)}(\lambda) = \frac{1}{2}(2\sigma^{(i)})^{-d}$, we have

$$E_{p_n}\{W_{\Lambda_n}^{(i)}\} \ge C_1^{(i)}(\lambda) s(n)\,. \tag{5.44}$$

This completes the proof of the lower bound. $\square$

Proof of Theorem 3.3. For this proof use (4.1), (4.2) and Postulates (III) and (IV). We start with a lower bound on $Pr_{p_n}\{W^{(i)}_{\Lambda_n} \geq K^{-1}E_{p_n}(W^{(i)}_{\Lambda_n})\}$. We again have (5.39) for some λ and n_1, and by Theorem 3.1 i) (whose proof only used (4.1) and Postulates (III) and (IV)) there exists some constant $C_2^{(i)}$, which depends on the sequence $\{p_n\}$, such that

$$E_{p_n}\{W^{(i)}_{\Lambda_n}\} \leq C_2^{(i)} s(n).$$

Thus if m is such that

$$s(m) \geq K^{-1} C_2^{(i)} s(n), \tag{5.45}$$

then

$$Pr_{p_n}\Big\{W^{(i)}_{\Lambda_n} \geq K^{-1}E_{p_n}\{W^{(i)}_{\Lambda_n}\}\Big\} \geq Pr_{p_n}\Big\{W^{(i)}_{\Lambda_n} \geq s(m)\Big\}. \tag{5.46}$$

We now choose $m = \lfloor n/\sigma^{(i)}(\lambda,\delta)\rfloor$, where the $\sigma^{(i)}$ are the constants introduced above (5.40). Then (5.45) will be satisfied for large enough K (by (5.43)). Since $n \geq n_1$ and $n \geq \sigma^{(i)}(\lambda,\delta)$ implies $m \geq 1$ and $m\sigma^{(i)}(\lambda,\delta) \leq n \leq \lambda L_0(p_n)$, we now can use (5.40) to conclude that

$$\liminf_{n\to\infty} Pr_{p_n}\Big\{W^{(i)}_{\Lambda_n} \geq K^{-1}E_{p_n}\{W^{(i)}_{\Lambda_n}\}\Big\} \geq 1-\delta\,, \tag{5.47}$$

provided K is large enough. Together with Markov's inequality,

$$Pr_{p_n}\{W^{(i)}_{\Lambda_n} \geq KE_{p_n}\{W^{(i)}_{\Lambda_n}\}\} \leq K^{-1}\,, \tag{5.48}$$

(5.47) implies Theorem 3.3 i).

In order to complete the proof of Theorem 3.3, we choose $m(n)$ as the maximal $m \leq (\sigma_4/\lambda \wedge 1)n$ such that $K^{-1}C_1^{(i)}(\lambda)s(n) > s(m)$, where σ_4 is as in Proposition 5.4, λ as in (5.39) and $C_1^{(i)}$ as in (5.44). Then, by (5.44) and $W^{(i)}_\Lambda \leq W^{(1)}_\Lambda$, we have

$$\begin{aligned}\limsup_{n\to\infty} Pr_{p_n}\Big\{W^{(i)}_{\Lambda_n} \geq K^{-1}E_{p_n}\{W^{(i)}_{\Lambda_n}\}\Big\} &\leq \limsup_{n\to\infty} Pr_{p_n}\Big\{W^{(i)}_{\Lambda_n} \geq K^{-1}C_1^{(i)}(\lambda)s(n)\Big\}\\ &\leq \limsup_{n\to\infty} Pr_{p_n}\Big\{W^{(1)}_{\Lambda_n} \geq K^{-1}C_1^{(i)}(\lambda)s(n)\Big\}\\ &\leq \limsup_{n\to\infty}\Big[1 - Pr_{p_n}\Big\{W^{(1)}_{\Lambda_n} \leq s(m(n))\Big\}\Big].\end{aligned} \tag{5.49}$$

Since $n/m(n)$ is bounded above by virtue of Postulate (IV) (see (5.10)), Proposition 5.4 shows that the right-hand side of (5.49) is bounded away from 1. This proves Theorem 3.3 ii). □

Proof of Theorem 3.4. For this proof we only use (4.1), and Postulates (III) and (IV). Theorem 3.4 follows immediately from Proposition 5.1. Indeed, let λ and n_1 be the constants from (5.39), and C_1, C_2 and σ_1 be those from Proposition 5.1. Choose $\sigma_2 \geq \max\{\sigma_1, \lambda\sigma_1, n_1\}$. We note that then $m \geq 1$ and $\sigma_2 m \leq n$ imply $n \geq n_1$, and hence $n \leq \lambda L_0(p_n)$ and $\sigma_1 m \leq L_0(p_n)$. The conditions of Theorem 3.4 therefore imply those of Proposition 5.1, proving that Theorem 3.4 under the assumption that (4.1), as well as Postulate (III) and (IV) hold. □

5.2 Above the Scaling Window.

In this subsection, we prove Theorem 3.1 iii) and Theorem 3.2. To this end, we consider separately those clusters $\mathcal{C}_\Lambda^{(i)}$ which intersect the infinite cluster $\mathcal{C}_\infty$ and those which do not. We denote the clusters intersecting $\mathcal{C}_\infty$ by $\mathcal{C}_{\Lambda,\infty}^{(1)}, \mathcal{C}_{\Lambda,\infty}^{(2)}, \cdots \mathcal{C}_{\Lambda,\infty}^{(k)}$, ordering them again from largest to smallest size, with lexicographic order between clusters of the same size. In the same way, $\mathcal{C}_{\Lambda,\mathrm{fin}}^{(1)}, \mathcal{C}_{\Lambda,\mathrm{fin}}^{(2)}, \cdots \mathcal{C}_{\Lambda,\mathrm{fin}}^{(k)}$ denote the clusters in Λ which do not intersect the infinite cluster $\mathcal{C}_\infty$. Finally, $W_{\Lambda,\mathrm{fin}}^{(i)} = |\mathcal{C}_{\Lambda,\mathrm{fin}}^{(i)}|$ and $W_{\Lambda,\infty}^{(i)} = |\mathcal{C}_{\Lambda,\infty}^{(i)}|$ denote the sizes of the ith largest clusters in the corresponding classes.

Proposition 5.5. *Suppose that Postulates (V) and (VI) hold. Then there exists a constant $C_1 < \infty$ such that*

$$\frac{E_p\{W_{\Lambda_n,\mathrm{fin}}^{(1)}\}}{|\Lambda_n| P_\infty(p)} \leq C_1 \Big(\frac{L_0(p)}{n}\Big)^{d/2} \quad \text{if} \quad P > p_c, \tag{5.50}$$

so that in particular

$$\frac{E_{p_n}\{W_{\Lambda_n,\mathrm{fin}}^{(1)}\}}{|\Lambda_n| P_\infty(p_n)} \to 0 \qquad \text{as } n \to \infty \tag{5.51}$$

whenever $p_n > p_c$ is a sequence of densities such that $n/L_0(p_n) \to \infty$ as $n \to \infty$.

Proof. Let $t(n) = (2nL_0(p))^{d/2} \pi_{L_0(p)}(p_c)$. Analogously to (5.2) we have

$$\begin{aligned} E_p\{W_{\Lambda_n,\mathrm{fin}}^{(1)}\} &\leq t(n) + E_p\{W_{\Lambda_n,\mathrm{fin}}^{(1)}; W_{\Lambda_n,\mathrm{fin}}^{(1)} \geq t(n)\} \\ &\leq t(n) + \sum_{v \in \Lambda_n} Pr_p\{|\mathcal{C}_{\Lambda_n}(v)| = W_{\Lambda_n,\mathrm{fin}}^{(1)}, |\mathcal{C}_{\Lambda_n}(v)| \geq t(n), v \not\leftrightarrow \infty\} \\ &\leq t(n) + |\Lambda_n| Pr_p\{|\mathcal{C}(\mathbf{0})| \geq t(n), \mathbf{0} \not\leftrightarrow \infty\}. \end{aligned} \tag{5.52}$$

Using Markov's inequality and Postulate (V) we obtain

$$\begin{aligned} E_p\{W_{\Lambda_n,\mathrm{fin}}^{(1)}\} &\leq t(n) + \frac{(2n)^d}{t(n)} \chi^{\mathrm{fin}}(p) \\ &\leq t(n) + D_4 \frac{(2n)^d}{t(n)} L_0^d(p) \pi_{L_0(p)}^2(p_c) \\ &= t(n)\big(1 + D_4\big)\,. \end{aligned} \tag{5.53}$$

Observing that $t(n)/|\Lambda_n| P_\infty(p) \asymp (L_0(p)/n)^{d/2}$ by Postulate (VI), we obtain (5.50) and hence (5.51). ☐

In order to estimate the size of the clusters $\mathcal{C}_{\Lambda,\infty}^{(1)}, \mathcal{C}_{\Lambda,\infty}^{(2)}, \cdots \mathcal{C}_{\Lambda,\infty}^{(k)}$, we make extensive use of the facts that

$$\sum_{i \geq 1} W_{\Lambda,\infty}^{(i)} = |\Lambda_n \cap \mathcal{C}_\infty| = \sum_{v \in \Lambda_n} I[v \leftrightarrow \infty] \tag{5.54}$$

and

$$E_{p_n}\{|\Lambda_n \cap \mathcal{C}_\infty|\} = \sum_{v \in \Lambda_n} Pr_{p_n}\{v \leftrightarrow \infty\} = |\Lambda_n| P_\infty(p_n)\,. \tag{5.55}$$

Lemma 5.6. *Suppose that Postulates (V) and (VI) hold. Let $p_n > p_c$ be a sequence of densities such that $n/L_0(p_n) \to \infty$ as $n \to \infty$. Then as $n \to \infty$*

$$\frac{|\Lambda_n \cap \mathcal{C}_\infty|}{|\Lambda_n| P_\infty(p_n)} \to 1 \qquad \textit{in probability}. \tag{5.56}$$

Proof. We bound the variance of $|\Lambda_n \cap \mathcal{C}_\infty|$ by

$$\begin{aligned} \mathrm{Var}_{p_n}\{|\Lambda_n \cap \mathcal{C}_\infty|\} &= \sum_{v,w\in\Lambda_n} \mathrm{Cov}_{p_n}(v \leftrightarrow \infty; w \leftrightarrow \infty) \\ &\le \sum_{\substack{v\in\Lambda_n \\ w\in\mathbb{Z}^d}} \mathrm{Cov}_{p_n}(v \leftrightarrow \infty; w \leftrightarrow \infty) = |\Lambda_n| \chi^{\mathrm{cov}}(p_n)\,. \end{aligned} \tag{5.57}$$

Note that we used here the positivity of $\mathrm{Cov}_{p_n}(v \leftrightarrow; w \leftrightarrow \infty)$; this follows from the Harris–FKG inequality. Combined with (5.55) and Postulates (V) and (VI), we obtain that for a suitable constant $C_1 < \infty$

$$\frac{\mathrm{Var}_{p_n}\{|\Lambda_n \cap \mathcal{C}_\infty|\}}{E^2_{p_n}\{|\Lambda_n \cap \mathcal{C}_\infty|\}} \le \frac{C_1 L_0(p_n)^d}{|\Lambda_n|} = C_1 \left(\frac{L_0(p_n)}{2n}\right)^d. \tag{5.58}$$

By our assumption on p_n, the right-hand side goes to zero as $n \to \infty$. This implies (5.56). ☐

Proposition 5.7. *Suppose that Postulates (II), (V) and (VI) hold. Let $p_n > p_c$ be a sequence of densities such that $n/L_0(p_n) \to \infty$ as $n \to \infty$. Then, as $n \to \infty$,*

$$\frac{W^{(1)}_{\Lambda_n,\infty}}{|\Lambda_n| P_\infty(p_n)} \to 1 \qquad \textit{in probability}. \tag{5.59}$$

Proof. We have to show that for all $\delta > 0$

$$Pr_{p_n}\{W^{(1)}_{\Lambda_n,\infty} \ge (1-\delta)|\Lambda_n| P_\infty(p_n)\} \to 1 \quad \text{as } n \to \infty \tag{5.60}$$

and

$$Pr_{p_n}\{W^{(1)}_{\Lambda_n,\infty} \le (1+\delta)|\Lambda_n| P_\infty(p_n)\} \to 1 \quad \text{as } n \to \infty. \tag{5.61}$$

Since $W^{(1)}_{\Lambda_n,\infty} \le |\Lambda_n \cap \mathcal{C}_\infty|$ by (5.54), the result (5.61) follows from (5.56). We are therefore left with proving (5.60). Again by (5.56), this amounts to showing that with high probability, the main contribution to the left-hand side of (5.54) comes from $W^{(1)}_{\Lambda_n,\infty}$.

We consider suitable volumes $\Lambda_m \subset \Lambda_n$ with

$$\lim_{n\to\infty} |\Lambda_m|/|\Lambda_n| > 1 - \delta\,. \tag{5.62}$$

Since

$$\frac{|\mathcal{C}_\infty \cap \Lambda_m|}{|\Lambda_m| P_\infty(p_n)} \to 1 \quad \text{in } Pr_{p_n}\text{-probability} \tag{5.63}$$

as $n \to \infty$ (the proof is identical to the proof of Lemma 5.6), we conclude that

$$Pr_{p_n}\{|\mathcal{C}_\infty \cap \Lambda_m| \geq (1-\delta)|\Lambda_n| P_\infty(p_n)\} \to 1 \quad \text{as} \quad n \to \infty\,. \tag{5.64}$$

We shall next show that for a suitable choice of Λ_m

$$P_{p_n}\Big\{\#\{i \mid \mathcal{C}^{(i)}_{\Lambda_n,\infty} \cap \Lambda_m \neq \emptyset\} \geq 2\Big\} \to 0 \quad \text{as} \quad n \to \infty\,. \tag{5.65}$$

If $\#\{i \mid \mathcal{C}^{(i)}_{\Lambda_n,\infty} \cap \Lambda_m \neq \emptyset\} = 1$, then all pieces of $\mathcal{C}_\infty \cap \Lambda_m$ are connected in Λ_n and

$$|\mathcal{C}^{(1)}_{\Lambda_n,\infty}| \geq |\mathcal{C}_\infty \cap \Lambda_m|,$$

so that (5.65) together with (5.64) will prove the desired result (5.60).

In order to show that Λ_m can be chosen so that (5.62) and (5.65) hold, we define, for $0 < \alpha < 1/6$ and $n \geq 1/\alpha$,

$$x = \frac{2}{\alpha} - 3, \quad L(n) = \lfloor \alpha n \rfloor, \quad M(n) = \lfloor x L(n) \rfloor \quad \text{and} \quad m = \frac{M(n)+1}{2}\,. \tag{5.66}$$

Note that with this choice $m < n$ for all $n \geq 1/\alpha$, and

$$\lim_{n\to\infty} \frac{|\Lambda_m|}{|\Lambda_n|} = \left(1 - \frac{3\alpha}{2}\right)^d\,. \tag{5.67}$$

A sufficiently small choice of α therefore ensures the condition (5.62). Note also that Λ_m is isomorphic to $[0, M(n)]^d$, while Λ_n is isomorphic to $[-\tilde{L}(n), M(n) + \tilde{L}(n)]^d$, where

$$\tilde{L}(n) := n - m \geq L(n)\,. \tag{5.68}$$

Using these observations and recalling the definition (2.23) of $S^{\text{fin}}_{\tilde{L},M}(p_n)$, we then bound

$$P_{p_n}\{\#\{i \mid \mathcal{C}^{(i)}_{\Lambda_n,\infty} \cap \Lambda_m \neq \emptyset\} \geq 2\} \leq S^{\text{fin}}_{\tilde{L}(n),M(n)}(p_n) \leq S^{\text{fin}}_{L(n),M(n)}(p_n)\,, \tag{5.69}$$

where in the last step we have used that $S^{\text{fin}}_{L,M}(p_n)$ is decreasing in L.

In order to complete the proof, we use that for any $\tilde{\varepsilon} > 0$,

$$L_0(p, \tilde{\varepsilon}; x) \asymp L_0(p) = L_0(p, \varepsilon; 1) \tag{5.70}$$

by Postulate (II). Our assumption $n/L_0(p_n) \to \infty$ therefore implies that $L_0(p_n, \tilde{\varepsilon}; x)/n$, and hence $L_0(p_n, \tilde{\varepsilon}, x)/L(n)$, goes to zero as $n \to \infty$. Since this is true for all $\tilde{\varepsilon} > 0$, we can use the definition (2.25) of $L_0(p_n, \tilde{\varepsilon}, x)$ to conclude that

$$S^{\text{fin}}_{L(n),M(n)}(p_n) = S^{\text{fin}}_{L(n),\lfloor xL(n)\rfloor}(p_n) \to 0 \quad \text{as} \quad n \to \infty\,. \tag{5.71}$$

Equations (5.71) and (5.69) imply (5.65), and hence the proposition. □

Proof of Theorem 3.1 iii). For this proof use Postulates (II) and (V) and (VI). Let $p_n > p_c$ be such that $n/L_0(p_n) \to \infty$. We may then use (5.59) to conclude that that

$$\liminf_{n\to\infty} \frac{E_{p_n}\{W^{(1)}_{\Lambda_n,\infty}\}}{|\Lambda_n| P_\infty(p_n)} \geq 1\,. \tag{5.72}$$

Since

$$E_{p_n}\{W^{(1)}_{\Lambda_n,\infty}\} \leq \sum_{i\geq 1} E_{p_n}\{W^{(i)}_{\Lambda_n,\infty}\} = |\Lambda_n| P_\infty(p_n) \tag{5.73}$$

for all n, it follows that

$$\lim_{n\to\infty} \frac{E_{p_n}\{W^{(1)}_{\Lambda_n,\infty}\}}{|\Lambda_n| P_\infty(p_n)} = 1\,. \tag{5.74}$$

Combined with (5.51) and $W^{(1)}_{\Lambda_n,\infty} \leq W^{(1)}_{\Lambda_n} \leq W^{(1)}_{\Lambda_n,\infty} + W^{(1)}_{\Lambda_n,\mathrm{fin}}$, this proves (3.10).

In order to prove (3.11), we note that (5.74) together with (5.54) and (5.72) imply that

$$\frac{E_{p_n}\{W^{(2)}_{\Lambda_n,\infty}\}}{|\Lambda_n| P_\infty(p_n)} \leq 1 - \frac{E_{p_n}\{W^{(1)}_{\Lambda_n,\infty}\}}{|\Lambda_n| P_\infty(p_n)} \to 0 \tag{5.75}$$

as $n \to \infty$. Combined with (5.51), this implies (3.11). □

Proof of Theorem 3.2. We again use Postulates (II) and (V) and (VI). As before, by assumption, $p_n > p_c$ for all sufficiently large n, and $n/L_0(p_n) \to \infty$. Using Markov's inequality and Proposition 5.5, we therefore get

$$\frac{W^{(1)}_{\Lambda_n,\mathrm{fin}}}{|\Lambda_n| P_\infty(p_n)} \to 0 \quad \text{in probability}\,. \tag{5.76}$$

Combined with Proposition 5.7, this implies Theorem 3.2. □

5.3 Below the Scaling Window.

We start with a lemma which will play a similar role below the window to that played by the lower bound in Proposition 5.1 inside the window.

Lemma 5.8. *Assume that (4.1) holds for some $\varepsilon > 0$ and that Postulates (III), (IV) and (VII) hold. Then there exist constants $0 < C_3 < \infty$ and $1 \leq \sigma_6, \sigma_7, \sigma_8 < \infty$ such that*

$$E_p\big\{\widetilde{N}_{\Lambda_n}(ks(L_0(p)), k\sigma_6 s(L_0(p)))\big\} \geq C_3 \frac{e^{-D_6 k}}{k} \Big(\frac{n}{L_0(p)}\Big)^d \tag{5.77}$$

provided $k \geq \sigma_7$, $n \geq \sigma_8 k L_0(p)$ and $p < p_c$. Here D_6 is the constant from Postulate (VII).

Proof. Let C_1 and C_2 be the constants from Lemma 4.4. Combining the bound (4.13) with Postulate (VII) and Proposition 4.6 we see that for suitable constants C_4, C_5, with $C_2 C_4 > D_6$, and k sufficiently large, say $k \geq C_7$, one gets

$$\begin{aligned} &Pr_p\{|\mathcal{C}(\mathbf{0})| \geq ks(L_0(p)),\ \text{but}\ \mathrm{diam}(\mathcal{C}(\mathbf{0})) < C_4 ks L_0(p)\} \\ &\geq Pr_p\{|\mathcal{C}(\mathbf{0})| \geq ks(L_0(p))\} - Pr_p\{\mathrm{diam}(\mathcal{C}(\mathbf{0})) \geq C_4 k L_0(p)\} \\ &\geq C_5 \pi_{L_0(p)}(p_c) e^{-D_6 k}. \end{aligned} \tag{5.78}$$

We want to restrict $|\mathcal{C}(\mathbf{0})|$ further. For this we use Lemma 4.7, which tells us that

$$Pr_p\{|\mathcal{C}(\mathbf{0})| \geq \sigma_6 ks(L_0(p))\} \leq C_1\pi_{L_0(p)}(p_c)e^{-C_2\sigma_6 k}.$$

Therefore, if we take $\sigma_6 > D_6/C_2$, then for sufficiently large k, say $k \geq \widetilde{C}_7$,

$$\begin{aligned} &Pr_p\{ks(L_0(p)) \leq |\mathcal{C}(\mathbf{0})| \leq \sigma_6 ks(L_0(p)),\ \text{but}\ \text{diam}(\mathcal{C}(\mathbf{0})) < C_4kL_0(p)\} \\ &\geq \frac{1}{2}C_5\pi_{L_0(p)}(p_c)e^{-D_6k}. \end{aligned} \tag{5.79}$$

Now let $n \geq 2C_4kL_0(p)$, $\tilde{n} = n - \lfloor C_4kL_0(p)\rfloor$, $\Lambda = \Lambda_r$ and $\widetilde{\Lambda} = \Lambda_{\tilde{n}}$. Observe that if $v \in \widetilde{\Lambda}$ and $\text{diam}(\mathcal{C}(v)) \leq \lfloor C_4kL_0(p)\rfloor$, then $\mathcal{C}(v) \subset \Lambda$ and $\mathcal{C}_\Lambda(v) = \mathcal{C}(v)$. Using this observation we now find

$$\begin{aligned} &E_p\big\{\widetilde{N}_{\Lambda_r}(ks(L_0(p)), k\sigma_6 s(L_0(p)))\big\} \\ &\geq \sum_{v\in\widetilde{\Lambda}}\sum_{s=ks(L_0(p))}^{\sigma_6 ks(L_0(p))}\frac{1}{s}Pr_p\{|\mathcal{C}_\Lambda(v)| = s,\ \text{but}\ \text{diam}(\mathcal{C}(v)) < \lfloor C_4kL_0(p)\rfloor\} \\ &\geq \frac{1}{\sigma_6 ks(L_0(p))}\sum_{v\in B}Pr_p\{ks(L_0(p)) \leq |\mathcal{C}(\mathbf{0})| \leq C_6ks(L_0(p)), \\ &\qquad\qquad \text{but}\ \text{diam}(\mathcal{C}(\mathbf{0})) < \lfloor C_4kL_0(p)\rfloor\} \\ &\geq C_3\frac{(2n)^d}{ks(L_0(p))}\pi_{L_0(p)}(p_c)e^{-D_6k} = C_3\Big(\frac{n}{L_0(p)}\Big)^d k^{-1}e^{-D_6k}. \end{aligned}$$

Choosing $\sigma_7 = \max\{C_7, \tilde{C}_7\}$ and $\sigma_8 = 2C_4$, this proves the lemma. □

Proof of Theorem 3.1 ii) and Theorem 3.5. For the proof, we will need (4.1), and Postulates (III), (IV) and (VII). Assume that $p_n < p_c$ for sufficiently large n, and $n/L_0(p_n) \to \infty$. It follows from (4.8) that for $z \geq 0$ and n large

$$\begin{aligned} E_{p_n}\{W_{\Lambda_n}^{(1)}\} &\leq s(L_0(p_n))\log\left(\frac{n}{L_0(p_n)}\right) \\ &\quad\times\Big[z + \int_z^\infty Pr_{p_n}\Big\{W_{\Lambda_n}^{(1)} \geq ys(L_0(p_n))\log\left(\frac{n}{L_0(p_n)}\right)\Big\}dy\Big] \\ &\leq s(L_0(p_n))\log\left(\frac{n}{L_0(p_n)}\right) \\ &\quad\times\Big[z + C_1\int_z^\infty\left(\frac{n}{L_0(p_n)}\right)^{d-C_2z}\exp[-C_2(y-z)\log\left(\frac{n}{L_0(p_n)}\right)]dy\Big]. \end{aligned}$$

By choosing $C_2z = d$ we see that

$$E_{p_n}\{W_{\Lambda_n}^{(1)}\} \leq C_3 s(L_0(p_n))\log\left(\frac{n}{L_0(p_n)}\right) \tag{5.80}$$

for a suitable constant $C_3 < \infty$. This proves the upper bound for $E_{p_n}\{W^{(1)}_{\Lambda_n}\}$, where we have so far only used Postulate (IV).

The lower bound for $E_{p_n}\{W^{(1)}_{\Lambda_n}\}$ follows immediately from Theorem 3.5 so that it suffices to prove (3.16). Also, we only have to prove that

$$\liminf_{n\to\infty} Pr_{p_n}\{K^{-1} \le \frac{W^{(i)}_{\Lambda_n}}{s(L_0(p_n))\log\frac{n}{L_0(p_n)}}\} \to 1, \text{ as } K \to \infty, \tag{5.81}$$

since the other part of (3.16) is obvious from Markov's inequality and the upper bound (5.80).

For brevity we write p instead of p_n for the remainder of this proof. The lower bound (5.77) will play a similar role to that played by Proposition 5.1. However, instead of using an analogue of Proposition 5.2 for a second moment, we now appeal to the BK-inequality [BK85]. This tells us that

$$\begin{aligned}&Pr_p\{\exists\, r \text{ disjoint clusters in } B_{\lceil\sigma_8 kL_0(p)\rceil} \text{ of size } \ge ks(L_0(p))\}\\ &\le \left[Pr_p\{\exists \text{ at least one cluster in } B_{\lceil\sigma_8 kL_0(p)\rceil} \text{ of size } \ge ks(L_0(p))\}\right]^r.\end{aligned}$$

Consequently if we set

$$\kappa = Pr_p\{\exists \text{ at least one cluster in } B_{\lceil\sigma_8 kL_0(p)\rceil} \text{ of size } \ge ks(L_0(p))\}, \tag{5.82}$$

then

$$E_p\{\text{number of disjoint clusters in } B_{\lceil\sigma_8 kL_0(p)\rceil} \text{ of size } \ge ks(L_0(p))\} \le \frac{\kappa}{1-\kappa}.$$

By (5.77) the left hand side here is at least $C_8 k^{d-1}\exp[-D_6 k]$, $C_8 = C_3\sigma_8^d$, so that

$$\kappa \ge \min\left(\frac{1}{2}, \frac{C_8}{2}k^{d-1}e^{-D_6 k}\right). \tag{5.83}$$

Now choose

$$k = k(n) = \left\lfloor C_9 \log\left(\frac{n}{L_0(p)}\right)\right\rfloor$$

with the constant $C_9 > 0$ but so small that $D_6 C_9 < d/2$. Then we can find in Λ_n approximately

$$\left(\frac{n}{2\sigma_8 kL_0(p)}\right)^d \ge \left(\frac{n}{L_0(p)}\right)^{d/2}$$

disjoint boxes $B_{\lceil\sigma_8 kL_0(p)\rceil}(v_i)$. Each of these boxes contains a cluster of size

$$\ge k(n)s(L_0(p)) \sim C_9 s(L_0(p_n))\log\left(\frac{n}{L_0(p_n)}\right) \tag{5.84}$$

with a probability at least

$$\min\left(\frac{1}{2}, \frac{C_8}{2} k^{d-1} e^{-D_6 k}\right). \tag{5.85}$$

Moreover, as in (4.16) we also have

$$Pr_p\{\partial B_{\lceil \sigma_8 k L_0(p)\rceil}(v_i) \not\leftrightarrow \partial B_{3\lceil \sigma_8 k L_0(p)\rceil}(v_i)\} \geq \varepsilon^{2d}.$$

For large n this gives

$$\begin{aligned} &Pr_p\{B_{\lceil \sigma_8 k L_0(p)\rceil}(v_i) \text{ contains a cluster of size } \frac{C_9}{2} s(L_0(p_n)) \log\left(\frac{n}{L_0(p_n)}\right) \\ &\qquad \text{and this cluster is not connected to } \partial B_{3\lceil \sigma_8 k L_0(p)\rceil}(v_i)\} \\ &\geq \tfrac{1}{2}\varepsilon^{2d} C_8 k^{d-1} \exp[-D_6 k]. \end{aligned} \tag{5.86}$$

Since the number of boxes times the right hand side of (5.86) tends to infinity (by our choice of $k(n)$ or C_9), the probability that at least i of these boxes contains a cluster of size (5.84), and that these clusters are not connected to each other tends to 1. This establishes (3.16). □

6. Verification of the Postulates in Two Dimensions

In this section we prove Theorem 3.6, which states that the Scaling Postulates (I) – (VII) hold for $d = 2$. Before we start on the proof we discuss some general tools. The fundamental tool for two-dimensional bond percolation is *duality.*[3] This rests on the following observations. Let $\mathbb{Z}^*$ denote the lattice $\mathbb{Z}^2 + (\frac{1}{2}, \frac{1}{2})$, which is called the dual lattice of $\mathbb{Z}^2$. Each dual edge e^* bisects exactly one edge e of the original lattice and vice versa. We call such a pair e^* and e, *associated.* For each configuration ω of occupied and vacant edges of $\mathbb{Z}^2$ we obtain a configuration on $\mathbb{Z}^*$ by declaring a dual edge e^* occupied (repectively, vacant) if its associated edge is occupied (respectively, vacant). It is a well known result that there exists an occupied horizontal crossing of the rectangle $[0, L] \times [0, M]$ if and only if there does not exist a vertical vacant dual crossing of $[\frac{1}{2}, L - \frac{1}{2}] \times [-\frac{1}{2}, M + \frac{1}{2}]$ (see [SmW78], Section 2.1 and [Kes82] , Sections 2.6, 2.4). This translates into

$$R_{L,M}(p) = 1 - R_{M+1,L-1}(1-p). \tag{6.1}$$

This relation can be used to relate quantities in the subcritical regime to similar quantities in the supercritical regime. For instance, define the two-dimensional finite-size scaling length as

$$\widetilde{L}_0(p, \varepsilon) = \begin{cases} \min\{L \mid R_{L,L}(p) \leq \varepsilon\} & \text{if } p < p_c \\ \min\{L \mid R_{L,L}(p) \geq 1 - \varepsilon\} & \text{if } p > p_c\,. \end{cases} \tag{6.2}$$

[3]Here we can use duality since we are dealing with bond percolation, which is self-dual. However, with a good deal more work, similar results can be proven for other two-dimensional models which are not self-dual – see [Kes87] (equation (1.23) and Section 4).

(Note that this is in the spirit of definition (1.21) of [Kes87]. However, [Kes87] treats bond percolation on $\mathbb{Z}^2$ as site percolation on the covering graph of $\mathbb{Z}^2$, so that the formal definition there is somewhat different. For the purposes of the proofs here this difference in the definitions is without significance.) It follows easily from duality and monotonicity of $R_{L,M}$ in L and M that for bond percolation on $\mathbb{Z}^2$, $\widetilde{L}_0(p,\varepsilon) \geq \widetilde{L}_0(1-p,\varepsilon)$ for $p < p_c$. From the rescaling lemma (Lemmas 3.4 and 4.12 in [ACCFR83]) and duality one obtains that for sufficiently small $\varepsilon > 0$ also $2\widetilde{L}_0(1-p,\varepsilon) - 1 \geq \widetilde{L}_0(p,\varepsilon)$ for $p < p_c$. We therefore have that

$$\widetilde{L}_0(p,\varepsilon) \asymp \widetilde{L}_0(1-p,\varepsilon), \quad p < p_c = \frac{1}{2}. \tag{6.3}$$

Similarly, using the rescaling lemma and the Russo-Seymour-Welsh lemma ([Rus78], [SW78], Section 3.4) it is straightforward to show that in $d = 2$, the definition (6.2) is equivalent to our finite-size scaling correlation length below threshold, see (2.20):

$$\widetilde{L}_0(p) \asymp L_0(p) \qquad \text{for} \quad p < p_c\,, \tag{6.4}$$

and to our finite-size scaling inverse surface tension above threshold, see (2.26):

$$\widetilde{L}_0(p) \asymp A_0(p) \qquad \text{for} \quad p > p_c\,. \tag{6.5}$$

As usual, the constants implicit in the equivalences (6.3)–(6.5) depend on ε.

It also follows from the Russo-Seymour-Welsh lemma that for each $x > 0$ and integer $k \geq 1$ there exists a constant $h(x,k,\varepsilon) > 0$ such that for $p \leq p_c$, $L \leq k\widetilde{L}_0(p)$, $M/L \geq x$, it holds that

$$R_{L,M}(p) \geq h(x,k,\varepsilon). \tag{6.6}$$

Thus, sponge crossing probabilities of rectangles with the ratio of the sides bounded away from 0 and ∞ and with a size comparable to $\widetilde{L}_0(p)$ are bounded away from 0. By means of the Harris-FKG inequality it is then also easy to see that the probability of an occupied circuit surrounding the origin in the annulus $A = [-M,M]^2 \setminus (-L,L)^2$ is bounded away from 0, provided $L \leq kL_0(p), M/L \geq 1 + x > 1$. Indeed, by obvious monotonicity we may assume that $M \leq 2L$. The annulus A is the union of four $M - L \times M$ rectangles, and if each of these has an occupied crossing in the long direction (i.e., a crossing in the direction of the side of length M), then A contains a circuit of the desired kind (compare [SmW78], Lemma 3.5). By the above, each of these crossings has a probability of $R_{M,M-L}(p) \geq h(x/(1+x) \wedge 1/2, 2k, \varepsilon)$, and by the Harris-FKG inequality the desired occupied circuit exists with a probability at least $h^4(x/(1+x) \wedge 1/2, 2k, \varepsilon)$.

Now consider two adjacent rectangles $[0,L]\times[0,M]$ and $[L,2L]\times[0,M]$, and assume that each of these contains an occupied horizontal crossing, r_1 and r_2, say. If, in addition there exist occupied vertical crossings of $[0,L]\times[-L,M+L]$ and $[L,2L]\times[-L,M+L]$ as well as occupied horizontal crossings of $[0,2L]\times[-L,0]$ and $[0,2L]\times[M,M+L]$, then these four crossings contain a circuit which necessarily intersects r_1 and r_2 and therefore connects r_1 and r_2 (see Figure 1). Therefore, another application of the Harris-FKG inequality shows that

$$
\begin{aligned}
&Pr_p\{\text{all horizontal crossings of } [0,L]\times[0,M] \text{ and of } [L,2L]\times[0,M] \text{ are connected}\big| \\
&\qquad\qquad \text{there exists at least one horizontal crossing in each of} \\
&\qquad\qquad [0,L]\times[0,M] \text{ and } [L,2L]\times[0,M]\} \\
&\geq h^2(\frac{L}{M+2L}, \frac{M+2L}{L_0(p)}, \varepsilon) h^2(\frac{1}{2}, \frac{2L}{L_0(p)}, \varepsilon). \qquad (6.7)
\end{aligned}
$$

If M/L and $L/L_0(p)$ are bounded, then the right hand side of (6.7) is bounded away from zero. By minor variations of this argument one sees that there is a lower bound for the probability that two occupied crossings r_1 and r_2 over length L which are within distance of order L from each other are connected (by a circuit of diameter also of order L), provided $L/L_0(p)$ is bounded. We shall say in such a situation that r_1 and r_2 can be connected by a *Harris ring.*

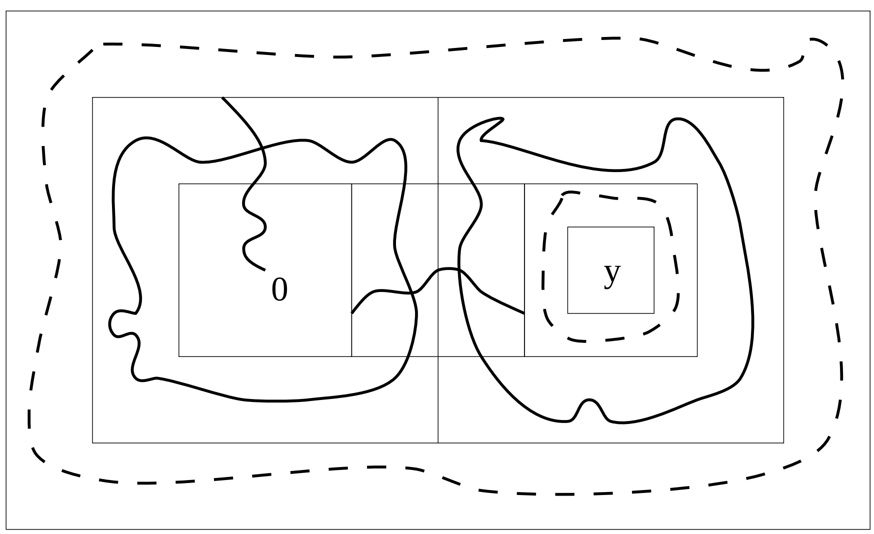

FIG. 1. Harris ring construction for the proof of (6.7)

We now prove the postulates for $d = 2$ in several subsections. These proofs rely to a large extent on the results and methods of [Kes86] and [Kes87].

6.1. Proof of Postulates (I) and (II).

Postulate (II) is the relation

$$A_0(p,\varepsilon) \asymp L_0(p,\varepsilon;1) \asymp L_0(p,\widetilde{\varepsilon};x) \qquad (6.8)$$

for all $p > p_c, x \geq 1$ and $\varepsilon, \widetilde{\varepsilon} \in (0, \varepsilon_0)$. Once we prove this, Postulate (I) follows e.g. from the equivalence in equation (6.5) and Postulate (II):

$$L_0(p) \asymp A_0(p) \asymp \widetilde{L}_0(p)\,, \qquad (6.9)$$

equation (6.3) and the known behavior (2.27). Hence it suffices to establish Postulate (II).

We claim that in order to prove (6.8), it suffices to show that for all $x \geq 1$ and $\widetilde{\varepsilon} \in (0, \varepsilon_0/2)$, there exists an $\varepsilon \in (0, \varepsilon_0)$ and a $\lambda = \lambda(\varepsilon, \widetilde{\varepsilon}, x)$ such that

$$\widetilde{L}_0(p, 2\varepsilon) \leq L_0(p, \widetilde{\varepsilon}, x) \leq \lambda \widetilde{L}_0(p, \varepsilon) + 1 \text{ as } p \downarrow p_c. \tag{6.10}$$

Indeed, given (6.10), we can deduce (6.8) for $\varepsilon, \widetilde{\varepsilon} < \varepsilon_0/2$ from (6.5) and the known equivalence of $\widetilde{L}_0(p, \varepsilon)$ at different values of ε, i.e.,

$$\widetilde{L}_0(p, \varepsilon_1) \asymp \widetilde{L}_0(p, \varepsilon_2) \text{ as } p \downarrow p_c \text{ for } 0 < \varepsilon_1, \varepsilon_2 < \varepsilon_0, \tag{6.11}$$

which follows from the rescaling lemma. Finally we must replace ε_0 by $\varepsilon_0/2$ to obtain Postulate (II).

We establish (6.10) via an upper and a lower bound. For the upper bound, we note that for all L, M

$$S^{\text{fin}}_{L,M}(p) \leq 1 - P_p(\exists \text{ an occupied circuit in } H_{L,M} \text{ surrounding } \partial_I H_{L,M}) \,. \tag{6.12}$$

Given $\varepsilon, \widetilde{\varepsilon} \in (0, \varepsilon_0)$ and $x \geq 1$, it is not hard to show, by means of the rescaling lemma (compare the argument for (6.7)), that there exists a $\lambda = \lambda(\varepsilon, \widetilde{\varepsilon}; x)$ such that if $M = \lfloor xL \rfloor$ and $L \geq \lambda \widetilde{L}_0(p, \varepsilon)$, then the probability of the circuit described in (6.12) is strictly bounded below by $1 - \widetilde{\varepsilon}$ for $p > p_c$. Hence $S^{\text{fin}}_{L,\lfloor xL \rfloor}(p) < \widetilde{\varepsilon}$ for all $L \geq \lambda \widetilde{L}_0(p, \varepsilon)$. But it follows from the definition (2.25) that $S^{\text{fin}}_{L,\lfloor xL \rfloor}(p) \geq \widetilde{\varepsilon}$ if $L = L_0(p, \widetilde{\varepsilon}; x) - 1$. Thus

$$L_0(p, \widetilde{\varepsilon}; x) \leq \lambda \widetilde{L}_0(p, \varepsilon) + 1 \,. \tag{6.13}$$

Next we establish a lower bound of the same form. To this end, note that the annulus $H_{L,\lfloor xL \rfloor}$ consists of four non-overlapping $L \times \lfloor xL \rfloor$ rectangles and four $L \times L$ corners. Let us call the rectangles the left, right, upper and lower rectangles. Clearly, for all L

$$\begin{aligned} S^{\text{fin}}_{L,\lfloor xL \rfloor}(p) \geq P_p(&\exists \text{ an occupied left-right crossing in the left rectangle and} \\ &\text{a vacant dual left-right crossing in the right rectangle,} \\ &\text{each connecting } \partial_I H_{L,\lfloor xL \rfloor} \text{ to } \partial_E H_{L,\lfloor xL \rfloor}) \,. \end{aligned} \tag{6.14}$$

Since $x \geq 1$, the lower bound in (6.14) is only strengthened by requiring that the occupied crossing occur in an $L \times L$ sub-box of the corresponding $L \times \lfloor xL \rfloor$ rectangle and that the vacant dual crossing occur in an $(L+1) \times (L-1)$ rectangle. By (6.1) this gives

$$S^{\text{fin}}_{L,\lfloor xL \rfloor}(p) \geq R_{L,L}(1 - R_{L,L}) \geq \frac{1}{2}(1 - R_{L,L}), \tag{6.15}$$

since $p > p_c$. Now let $L = L_0(p, \widetilde{\varepsilon}; x)$. It then follows from the definition (2.25) that $S^{\text{fin}}_{L,\lfloor xL \rfloor}(p) \leq \widetilde{\varepsilon}$, so that (6.15) implies $R_{L,L} \geq 1 - (2\widetilde{\varepsilon})$ if $L = L_0(p, \widetilde{\varepsilon}; x)$. Comparing this with the definition (6.2) for $p > p_c$, we conclude

$$L_0(p, \widetilde{\varepsilon}; x) \geq \widetilde{L}_0(p, 2\widetilde{\varepsilon}) \,, \tag{6.16}$$

a lower bound of the desired form. □

6.2. Proof of Postulate (III).

Postulate (III) is almost identical to Theorem 1 of [Kes87], except that the latter uses the condition $n \leq \widetilde{L}_0(p,\varepsilon)$, whereas Postulate (III) assumes $n \leq L_0(p,\varepsilon)$. Thus, to establish Postulate (III), it suffices to show that for all $\varepsilon \in (0,\varepsilon_0)$, there exists an $\widetilde{\varepsilon} \in (0,\varepsilon_0)$ such that

$$L_0(p,\varepsilon) \leq \widetilde{L}_0(p,\widetilde{\varepsilon})\,. \tag{6.17}$$

To prove (6.17) we note that by (6.9) we have $L_0(p,\varepsilon) \leq \lambda(\varepsilon)\widetilde{L}_0(p,\varepsilon)$ for $p > p_c$ and a suitable $\lambda < \infty$. This relation also holds for $p < p_c$, as observed in (6.4). Therefore, it suffices to show that for all $p \neq p_c$, $\lambda < \infty$, and $\varepsilon \in (0,\varepsilon_0)$, there exists an $\widetilde{\varepsilon} \in (0,\varepsilon_0)$ such that

$$\lambda\widetilde{L}_0(p,\varepsilon) \leq \widetilde{L}_0(p,\widetilde{\varepsilon})\,. \tag{6.18}$$

Finally, by (6.3), it suffices to establish (6.18) only for $p < p_c$, and by iteration, to establish the latter only for $\lambda = 2$. To this end, we note that by the Russo-Seymour-Welsh lemma ([Rus78],[SW78], Section 3.4), rescaling and the obvious monotonicity of $R_{L,M}$, we have

$$R_{M,M}(p) \geq f(R_{L,L}(p)) \text{ if } L \leq M \leq 3L\,, \tag{6.19}$$

for some function f on $[0,1]$ which is strictly positive on $(0,1]$. Without loss of generality we may take $f(\varepsilon) \leq \varepsilon$. Using the definition (6.2) of $\widetilde{L}_0(\varepsilon,p)$, we conclude that

$$R_{M,M}(p) > f(\varepsilon) \qquad \text{if} \qquad M \leq 3\widetilde{L}_0(p,\varepsilon) - 3\,. \tag{6.20}$$

As a consequence,

$$\widetilde{L}_0(p,f(\varepsilon)) \geq 3\widetilde{L}_0(\varepsilon,p) - 2 \geq 2\widetilde{L}_0(p,\varepsilon)\,, \tag{6.21}$$

where we have used that $R_{1,1}(p) \geq p > \varepsilon$, and hence $\widetilde{L}_0(p,\varepsilon) > 1$ in the last step. This establishes (6.18) and hence Postulate (III). □

6.3. Proof of Postulate (IV).

We will establish Postulate (IV) for all p such that $m \leq L_0(p)$ (a somewhat stronger result than the stated postulate at p_c). This postulate with $\rho_1 = 2$ follows from the claim that for some $C_1 > 0$

$$\frac{\pi_m(p)}{\pi_n(p)} \geq C_1\Big(\frac{m}{n}\Big)^{-1/2} \quad \text{if} \quad n \leq m \leq L_0(p)\,. \tag{6.22}$$

In order to establish (6.22), we assume that $kn \leq m \leq (k+1)n$ for some integer $k \geq 1$. By (2.12) and monotonicity of π_n,

$$\pi_m \geq \pi_{(k+1)n} \geq \widetilde{\pi}_{(k+1)n}\,. \tag{6.23}$$

Recall the definition (2.9) of $\widetilde{\pi}_{(k+1)n}$ and observe that one mechanism to ensure that the origin is connected to the line at $x_1 = (k+1)n$ is to have (1) the origin connected to some point in $\partial B_n(0)$, (2) some point on $\partial B_n(0)$ connected to the line at $x_1 = (k+1)n$, and (3)

Harris rings in the annuli $B_n \setminus B_{n/2}$ and $B_{2n} \setminus B_n$ and a rectangle crossing from (say) the right boundary of $B_{n/2}$ to the central quarter of the right boundary of B_{2n} to "glue" the connections in (1) and (2) together. Since $n \le L_0(p)$, the probability of the third event is bounded away from zero, uniformly in n (as in (6.7)). Denote the probability of the event described in (2) above by $G_{n,kn}$. Equation (6.23) and the Harris-FKG inequality then imply that for some constant $C_2 > 0$

$$\pi_m \ge C_2 \pi_n G_{n,kn} \,. \tag{6.24}$$

By an argument almost identical to the proof of corollary (3.15) in [BK85], $kn \le L_0(p)$ implies $G_{n,kn} \ge C_3/\sqrt{k}$, where C_3 is a lower bound on the probability of an occupied crossing of a $2kn \times 2kn$ square. The constant $C_3 > 0$ by virtue of (6.17) and (6.20). (Essentially this same argument is used in [Kes87], equation (3.6) and its proof on p. 143.) Thus (6.24) implies the desired bound (6.22). □

6.4. Proof of Postulate (V).

Theorem 3 of [Kes87] gives the second inequality in Postulate (V). Thus it suffices to prove that for a suitable constant D_4 and all $p > p_c$

$$\chi^{\rm cov}(p) \le D_4 L_0^2(p) \pi^2_{L_0(p)}(p_c) \,. \tag{6.25}$$

To this end, we decompose the sum defining $\chi^{\rm cov}(p)$ (with $|v|$ short for $|v|_\infty$ and L_0 for $L_0(p)$):

$$\chi^{\rm cov}(p) = \sum_{|v| \le 2L_0} {\rm Cov}_p(\mathbf{0} \leftrightarrow \infty; v \leftrightarrow \infty) + \sum_{|v| > 2L_0} {\rm Cov}_p(\mathbf{0} \leftrightarrow \infty; v \leftrightarrow \infty) \,. \tag{6.26}$$

To control the first term, we use the bound (4.4) in Lemma 4.1 and Postulate III to estimate

$$\begin{aligned} \sum_{|v| \le 2L_0} {\rm Cov}_p(\mathbf{0} \leftrightarrow \infty; v \leftrightarrow \infty) &\le \sum_{|v| \le 2L_0} P_p\{\mathbf{0} \leftrightarrow \infty, v \leftrightarrow \infty\} \\ &\le \sum_{|v| \le 2L_0} \tau(\mathbf{0}, v) \le \sum_{|v| \le 2L_0} \pi^2_{[|v|/2]}(p) \asymp L_0^2 \pi^2_{L_0}(p_c) \,. \end{aligned} \tag{6.27}$$

Next, we bound the second term in (6.26). To this end, let $B(w) = B_{L_0}(w)$ be the box of radius L_0 centered at w. For $|v| > 2L_0$, we have

$$\begin{aligned} {\rm Cov}_p(\mathbf{0} \leftrightarrow \infty; v \leftrightarrow \infty) &= {\rm Cov}_p(\mathbf{0} \not\leftrightarrow \infty; v \not\leftrightarrow \infty) \\ &= {\rm Cov}_p(\mathbf{0} \not\leftrightarrow \infty, \mathbf{0} \leftrightarrow \partial B(\mathbf{0}); v \not\leftrightarrow \infty, v \leftrightarrow \partial B(v)) \\ &\quad + {\rm Cov}_p(\mathbf{0} \not\leftrightarrow \infty, \mathbf{0} \leftrightarrow \partial B(\mathbf{0}); v \not\leftrightarrow \partial B(v)) \\ &\quad + {\rm Cov}_p(\mathbf{0} \not\leftrightarrow \partial B(\mathbf{0}); v \not\leftrightarrow \infty, v \leftrightarrow \partial B(v)) \\ &\quad + {\rm Cov}_p(\mathbf{0} \not\leftrightarrow \partial B(\mathbf{0}); v \not\leftrightarrow \partial B(v)) \\ &= {\rm Cov}_p(\mathbf{0} \not\leftrightarrow \infty, \mathbf{0} \leftrightarrow \partial B(\mathbf{0}); v \not\leftrightarrow \infty, v \leftrightarrow \partial B(v)) \\ &\quad + 2\,{\rm Cov}_p(\mathbf{0} \not\leftrightarrow \partial B(\mathbf{0}); v \not\leftrightarrow \infty, v \leftrightarrow \partial B(v)) \,, \end{aligned} \tag{6.28}$$

where in the last step we have used that $\mathrm{Cov}_p(\mathbf{0} \not\leftrightarrow \partial B(\mathbf{0}); v \not\leftrightarrow \partial B(v)) = 0$ by the independence of the events $\{\mathbf{0} \not\leftrightarrow \partial B(\mathbf{0})\}$ and $\{v \not\leftrightarrow \partial B(v)\}$ when $B(\mathbf{0})$ and $B(v)$ are disjoint, and also the symmetry of the roles played by $\mathbf{0}$ and v. Now we bound the second term on the right hand side of (6.28) according to

$$\begin{aligned}
&\mathrm{Cov}_p(\mathbf{0} \not\leftrightarrow \partial B(\mathbf{0}); v \not\leftrightarrow \infty, v \leftrightarrow \partial B(v)) \\
&= \mathrm{Cov}_p(\mathbf{0} \not\leftrightarrow \partial B(\mathbf{0}); v \not\leftrightarrow \infty, v \leftrightarrow \partial B(v), v \leftrightarrow \partial B(\mathbf{0})) \\
&\qquad + \mathrm{Cov}_p(\mathbf{0} \not\leftrightarrow \partial B(\mathbf{0}); v \not\leftrightarrow \infty, v \leftrightarrow \partial B(v), v \not\leftrightarrow \partial B(\mathbf{0})) \\
&= \mathrm{Cov}_p(\mathbf{0} \not\leftrightarrow \partial B(\mathbf{0}); v \not\leftrightarrow \infty, v \leftrightarrow \partial B(\mathbf{0})) \\
&= -\mathrm{Cov}_p(\mathbf{0} \leftrightarrow \partial B(\mathbf{0}); v \not\leftrightarrow \infty, v \leftrightarrow \partial B(\mathbf{0})) \\
&\leq P_p\{\mathbf{0} \leftrightarrow \partial B(\mathbf{0})\}\, P_p\{v \not\leftrightarrow \infty, v \leftrightarrow \partial B(\mathbf{0})\}\,,
\end{aligned} \tag{6.29}$$

where we have used that the two events $\{v \not\leftrightarrow \infty\} \cap \{v \leftrightarrow \partial B(v)\} \cap \{v \not\leftrightarrow \partial B(\mathbf{0})\}$ and $\mathbf{0} \not\leftrightarrow \partial B(\mathbf{0})$ are independent. Using the Harris-FKG inequality and obvious monotonicities, the second factor on the right hand side of (6.29) is in turn bounded according to

$$\begin{aligned}
&Pr_p\{v \not\leftrightarrow \infty, v \leftrightarrow \partial B(\mathbf{0})\} \\
&\leq Pr_p\{v \leftrightarrow \partial B(\mathbf{0})\} \\
&\qquad Pr_p\{\exists w \in \partial B(\mathbf{0}) \text{ such that } w \text{ and } v \text{ are surrounded by a vacant dual contour}\}
\end{aligned} \tag{6.30}$$

We now follow a coarse-graining argument along the lines of the proof of Theorem 3 in [Kes87] (see (3.12), (3.13) and (2.25) there). Let $v = (v_1, v_2)$ and for the sake of argument let $v_1 = |v| = |v|_\infty$. If there exists a vacant dual contour surrounding $w \in \partial B(\mathbf{0})$ and v, then there exists a vacant dual path from $B(\mathbf{0})$ to some $B(v_1 + j, v_2)$ with $j \geq 0$. By (2.25) in [Kes87] the probability that such a vacant path exists is at most $C_1 \exp[-C_2|v|/L_0]$. Together with (6.29) and Postulate III this leads to a bound of

$$C_3 \pi_{L_0}^2(p_c) \exp[-C_2|v|/L_0] \tag{6.31}$$

for the second term in the right hand side of (6.28).

Next we bound the first term in the right hand side of (6.28) by means of the BK inequality as follows:

$$\begin{aligned}
&Pr_p\{\mathbf{0} \not\leftrightarrow \infty, \mathbf{0} \leftrightarrow \partial B(\mathbf{0}); v \not\leftrightarrow \infty, v \leftrightarrow \partial B(v)\} \\
&\leq Pr_p\{\mathbf{0} \leftrightarrow \partial B(\mathbf{0}), v \leftrightarrow \partial B(v) \text{ and there exist vacant dual contours} \\
&\qquad\quad \mathcal{C}_1, \mathcal{C}_2 \text{ surrounding } \mathbf{0} \text{ and } v, \text{ respectively}\} \\
&\leq Pr_p\{\mathbf{0} \leftrightarrow \partial B(\mathbf{0}), v \leftrightarrow \partial B(v) \text{ and there exist } \textit{edge-disjoint} \text{ vacant dual contours} \\
&\qquad\quad \mathcal{C}_1, \mathcal{C}_2 \text{ surrounding } \mathbf{0} \text{ and } v, \text{ respectively}\} \\
&\quad + Pr_p\{\mathbf{0} \leftrightarrow \partial B(\mathbf{0}), v \leftrightarrow \partial B(v) \text{ and there exist vacant dual contours } \mathcal{C}_1, \mathcal{C}_2 \\
&\qquad\quad \text{surrounding } \mathbf{0} \text{ and } v, \text{ respectively, and } C_1 \text{ and } C_2 \text{ have an edge in common}\}.
\end{aligned}$$

By the BK inequality the first term in the right hand side is no more than

$$\begin{aligned} &Pr_p\{\mathbf{0} \leftrightarrow \partial B(\mathbf{0}) \text{ and there exists a vacant dual contour } \mathcal{C}_1 \text{ which surrounds } \mathbf{0}\} \\ &\qquad \times Pr_p\{v \leftrightarrow \partial B(v) \text{ and there exists a vacant dual contour } \mathcal{C}_2 \text{ which surrounds } v\} \\ &= Pr_p\{\mathbf{0} \not\leftrightarrow \infty, \mathbf{0} \leftrightarrow \partial B(\mathbf{0})\} Pr_p\{v \not\leftrightarrow \infty, v \leftrightarrow \partial B(v)\}. \end{aligned}$$

Therefore,

$$\begin{aligned} &\mathrm{Cov}_p(\mathbf{0} \not\leftrightarrow \infty, \mathbf{0} \leftrightarrow \partial B(\mathbf{0}); v \not\leftrightarrow \infty, v \leftrightarrow \partial B(v)\} \\ &\le Pr_p\{\mathbf{0} \leftrightarrow \partial B(\mathbf{0}), v \leftrightarrow \partial B(v) \text{ and there exist vacant dual contours } \mathcal{C}_1, \mathcal{C}_2 \\ &\qquad \text{surrounding } \mathbf{0} \text{ and } v, \text{ respectively, and } C_1 \text{ and } C_2 \text{ have an edge in common}\} \\ &\le Pr_p\{\mathbf{0} \leftrightarrow \partial B(\mathbf{0})\} Pr_p\{v \leftrightarrow \partial B(v)\} Pr_p\{\exists \text{ vacant dual contours } \mathcal{C}_1, \mathcal{C}_2 \\ &\qquad \text{surrounding } \mathbf{0} \text{ and } v \text{ respectively, and } \mathcal{C}_1 \text{ and } \mathcal{C}_2 \text{ have an edge in common}\}, \end{aligned} \tag{6.32}$$

where we have used the Harris-FKG inequality and disjointness of $B(\mathbf{0})$ and $B(v)$ in the last step. If the two dual contours $\mathcal{C}_1, \mathcal{C}_2$ in (6.32) have an edge in common, and if again $v = (v_1, v_2)$ with $v_1 = |v|$, then $\mathcal{C}_1 \cup \mathcal{C}_2$ contains a vacant dual path from some $B(-j_1, 0)$ to some $B(v_1 + j_2, v_2)$ with $j_1, j_2 \ge 0$. The same argument as used for (6.31) now shows that also the first term in the right hand side of (6.28) is bounded by (6.31). Finally, then

$$\sum_{|v| > 2L_0} \mathrm{Cov}_p(\mathbf{0} \leftrightarrow \infty; v \leftrightarrow \infty) \le \sum_{|v| > 2L_0} 2C_3 \pi^2_{L_0}(p_c) \exp[-C_2 |v|/L_0] \le C(\varepsilon) L_0^2 \pi^2_{L_0}(p_c)\,. \tag{6.33}$$

Together with (6.26), (6.27) this yields (6.25). □

6.5. Proof of Postulate (VI).

Postulate (VI) for $d = 2$ goes back to [Ngu85]. We can also immediately obtain this from Theorem 2 in [Kes87], which states that $P_\infty(p)$ is of the same order as $\pi_{\widetilde{L}_0(p,\varepsilon)}(p_c)$. But by (6.10), (6.11) there exists a $\lambda = \lambda(\varepsilon) \ge 1$ such that $\widetilde{L}_0(p, \varepsilon) \le \lambda L_0(p, \varepsilon)$. Therefore, by Postulate (IV)

$$\pi_{\widetilde{L}_0(p,\varepsilon)}(p_c) \ge \pi_{\lambda L_0(p,\varepsilon)}(p_c) \ge D_3 \lambda^{-1/\rho_1} \pi_{L_0(p,\varepsilon)}(p_c).$$

Combined with Theorem 2 of [Kes87] this gives one inequality of Postulate (VI). In the other direction, it is trivial that $P_\infty(p) \le \pi_{L_0(p)}(p)$ and further $\pi_{L_0(p)}(p) \asymp \pi_{L_0(p)}(p_c)$, by Postulate (III). □

6.6. Proof of Postulate (VII).

We shall build a cluster of size at least $ks(L_0(p))$ by connecting together $C_1 k$ clusters of size at least $C_2 s(L_0(p))$ (for suitable constants C_1, C_2) in adjacent squares of size $2L_0(p)$. These clusters will be connected by means of Harris rings.

By Postulate (IV) and Proposition 4.6 (which relies only on Postulate (I) — (IV)) there exists a $\sigma_0 \in (0,1]$ such that for $n_0 = \lfloor \sigma_0 L_0(p)/2 \rfloor$

$$P_{\geq s(L_0(p))}(p) \leq P_{\geq s(n_0)}(p) \leq C_3 \pi_{n_0}(p_c);$$

in the first inequality we used that Postulate (IV) implies (5.10), which in turns implies that $s(m) \geq s(n)$ if m/n is large enough. In turn, by Postulate (III), the right hand side here is at most $C_4 \pi_{n_0}(p)$. It therefore suffices to show that for $p < p_c$ and suitable constants C_5, D_6

$$P_{\geq ks(L_0(p))}(p) \geq C_5 e^{-D_6 k} \pi_{n_0}(p). \tag{6.34}$$

First we use Theorem 3.3 and Lemma 4.4, (4.14). These results show that there exist constants $K_0 < \infty$ and $y_0 > 0$ such that

$$\begin{aligned} &Pr_p\{\exists \text{ cluster } \mathcal{C} \subset \Lambda_{L_0(p)} \text{ with } \mathrm{diam}(\mathcal{C}) \geq y_0 L_0(p) \text{ and } |\mathcal{C}| \geq K_0^{-1} s(L_0(p))\} \\ &= Pr_p\{W^{(1)}_{L_0(p)} \geq K_0^{-1} s(L_0(p))\} \\ &\qquad - Pr_p\{\exists \text{ cluster } \mathcal{C} \subset \Lambda_{L_0(p)} \text{ with } \mathrm{diam}(\mathcal{C}) < y_0 L_0(p) \text{ and } |\mathcal{C}| \geq K_0^{-1} s(L_0(p))\} \\ &\geq \frac{1}{2} - C_1 y_0^{-2} \exp[-C_2 (K_0 y_0)^{-1}] \geq \frac{1}{4}, \end{aligned} \tag{6.35}$$

provided $L_0(p) \geq 4/y_0$. The estimate (6.35) shows that with a probability of at least 1/4 there is a cluster with a "large" size and "large" diameter in $\Lambda_{L_0(p)}$. We wish to locate this large cluster more precisely. In fact we want to show that we may assume that it crosses a certain rectangle in the first coordinate direction. To this end we note that if $\mathrm{diam}(\mathcal{C}) \geq y_0 L_0(p)$, then there are two points $v, w \in \mathcal{C}$ so that $w_i - v_i \geq y_0 L_0(p)$ for $i = 1$ or $i = 2$. Assume that this holds for $i = 1$. Then for some $-2/y_0 \leq j \leq 2/y_0$ the event

$$\begin{aligned} M(p,j) :=& \{\exists \text{ cluster } \mathcal{C} \in \Lambda_{L_0(p)} \text{ with } |\mathcal{C}| \geq K_0^{-1} s(L_0(p)) \text{ that contains} \\ & \qquad \text{points } v, w \text{ with } v_1 \leq j y_0 L_0(p)/2 < (j+1) y_0 L_0(p)/2 \leq w_1\} \end{aligned} \tag{6.36}$$

must occur. Therefore there exists a $j_0 \in [-2/y_0, 2/y_0]$ for which

$$Pr_p\{M(p, j_0)\} \geq \frac{y_0}{8(y_0+1)}. \tag{6.37}$$

From (6.37) and translation invariance it follows that each of the events

$$\begin{aligned} &\{\exists \text{ cluster } \mathcal{C}_\ell \in [2\ell L_0(p), (2\ell+2)L_0(p)) \times [-L_0(p), L_0(p)) \text{ with } |\mathcal{C}_\ell| \geq K_0^{-1} s(L_0(p)) \\ &\qquad \text{and which crosses } [(2\ell + j_0) y_0 L_0(p)/2, (2\ell + j_0 + 1) y_0 L_0(p)/2] \times [-L_0(p), L_0(p)) \\ &\qquad \text{in the horizontal direction}\}\,, \ \ell \geq 0, \end{aligned} \tag{6.38}$$

has probability at least $y_0/(8y_0+8)$. Let $k \geq 1$ be given and take $r = \lceil kK_0 \rceil$. If the event in (6.38) occurs for $\ell = 0, 1, \dots, r$ and $\mathbf{0} \leftrightarrow \partial B_{n_0}(\mathbf{0})$, and the paths from $\mathbf{0}$ to $\partial B_{n_0}(\mathbf{0})$ and the horizontal crossings of $[(2\ell + j_0) y_0 L_0(p)/2, (2\ell + j_0 + 1) y_0 L_0(p)/2] \times [-L_0(p), L_0(p))\,, 0 \leq$

$\ell \leq r$ are all connected by Harris rings, then the cluster of the origin has size at least $rK_0^{-1}s(L_0(p)) \geq ks(L_0(p))$. The Harris-FKG inequality now shows that

$$P_{\geq ks(L_0(p))}(p) \geq \pi_{n_0}(p) C_6 [C_6 \frac{y_0}{8(y_0+1)}]^r.$$

This proves (6.34) with

$$D_6 = \log \frac{8(y_0+1)(K_0+1)}{C_6 y_0},$$

and Postulate (VII) follows for all $p < p_c$ with $L_0(p) \geq 4/y_0$. If $L_0(p) < 4/y_0$, the postulate follows from the trivial bound $P_{\geq ks(L_0(p))} \geq p^{ks(L_0(p))} \geq p^{64 y_0^{-2} k}$. □

7. Proof of Theorem 3.7.

In this section, we introduce Postulate (VII alt), which is slightly stronger than Postulate (VII), and prove Theorem 3.7. To state the Postulate (VII alt), we need some notation. For $k \geq 1$, let $[k]^d = \{1, \dots, k\}^d$. Given an integer $k \geq 1$, and a choice of vertices $v(\mathbf{j})$ in $\widetilde{\Lambda}(\mathbf{j}) := 2\mathbf{j}L_0(p) + \Lambda_{\lfloor L_0(p)/4 \rfloor}$, $\mathbf{j} \in [k]^d$, we define sets $\Lambda(\mathbf{j}) = 2\mathbf{j}L_0(p) + \Lambda_{L_0(p)}$ and

$$\Xi = \bigcup_{\mathbf{j} \in [k]^d} \Lambda(\mathbf{j}),$$

as well as events

$$G(\mathbf{j}) = G(\mathbf{j}; x) = \{|\mathcal{C}_{\Lambda(\mathbf{j})}(v(\mathbf{j}))| \geq xs(L_0(p))\},$$

$$G_k = G_k(x) = \bigcap_{\mathbf{j} \in [k]^d} G(\mathbf{j}; x),$$

$$H(\mathbf{j}) = \{v(\mathbf{j}) \leftrightarrow v(\mathbf{j} \pm e_i) \text{ in } \Xi,\ 1 \leq i \leq d\},$$

where the i-th component of $\mathbf{j} \pm e_i$ equals $j_i \pm 1$. We also define

$$H_k = \{\text{all } v(\mathbf{j}) \text{ with } \mathbf{j} \in [k]^d \text{ are connected in } \Xi\} = \bigcap_{\substack{2 \leq j_i \leq k-1 \\ 1 \leq i \leq d}} H(\mathbf{j}).$$

Postulate (VII alt) For all $0 < x \leq 1$ there exists a constant $D_7 = D_7(x) > 0$ such that

$$Pr_p\{H_k \mid G_k(x)\} \geq D_7^{k^d} \tag{7.1}$$

for all $\zeta_0 \leq p < p_c$, $k \geq 1$ and all choices of $v(\mathbf{j})$, $\mathbf{j} \in [k]^d$. We remind the reader that ζ_0 is some arbitrary number in $(0, p_c)$.

Note that there are k^d choices for $\mathbf{j} \in [k]^d$. Condition (7.1) therefore roughly speaking says that the conditional probability of $H(\mathbf{j})$, given that $|\mathcal{C}(v(\mathbf{j}))| \geq xs(L_0(p))$ and each $|\mathcal{C}(v(\mathbf{j} \pm e_i))| \geq xs(L_0(p))$, is at least D_7. Or still more intuitively, "clusters of size of order $s(L_0(p))$ and a distance of order $L_0(p)$ apart have a reasonable conditional probability of being connected." We also mention that (7.1) is actually not needed for all $x \in (0, 1]$, but only for one fixed value of x for which $Pr_p\{G(\mathbf{j})\} \geq C_1 \pi_{L_0(p)}(p_c)$ for some constant $C_1 > 0$, independent of $p < p_c$. Such x and C_1 can be shown to exist by means of the bound (5.40) which follows from Proposition 5.3.

7.1. Proof of Theorem 3.7i.

In this subsection we always assume Postulates (I)–(IV) and $\zeta_0 \le p < p_c$. For brevity we write in many places L for $L_0(p)$ and Λ for $\Lambda_{L_0(p)}$. In steps i–v we also use Postulate (VII alt), but we only shall use that (7.1) is valid for $0 < x \le x_0$ for some $x_0 > 0$. The value of x_0 is irrelevant. All constants in this section are independent of k.

Step i There exists an $x \in (0,1]$ and a constant $C_2 > 0$ such that uniformly for $v(\mathbf{j}) \in \widetilde{\Lambda}(\mathbf{j}) = 2\mathbf{j}L_0(p) + \Lambda_{\lfloor L_0(p)/4 \rfloor}$,

$$Pr_p\{G(\mathbf{j})\} \ge C_2 \pi_L(p_c). \tag{7.2}$$

To prove (7.2) we use the relation (5.3) between the distribution of $W_\Lambda^{(1)}$ and $P_{\ge s}$. For $r \ge 1$ and any $0 < C_1 < \infty$, we get

$$\begin{aligned} Pr_p\{W_{\Lambda_r}^{(1)} \ge C_1 s(r)\} &\le \sum_{v \in \Lambda_r} \sum_{s \ge C_1 s(r)} \frac{1}{s} Pr_p\{|\mathcal{C}_{\Lambda_r}(v)| = s\} \\ &\le \frac{|\Lambda_r|}{C_1 s(r)} \sup_{v \in \Lambda_r} Pr_p\{|\mathcal{C}_{\Lambda_r}(v)| \ge C_1 s(r)\}. \end{aligned} \tag{7.3}$$

On the other hand, by (5.10), $s(m) \ge s(r)$ and hence

$$Pr_p\{W_{\Lambda_r}^{(1)} \ge C_1 s(r)\} \ge Pr_p\{W_{\Lambda_r}^{(1)} \ge s(m)\} \tag{7.4}$$

whenever $m \ge r(C_1/D_3)^{2/d}$. Setting $r = \lfloor L_0(p)/4 \rfloor$, $m = r\lceil (C_1/D_3)^{2/d} \rceil$, and choosing $C_1 > 0$ small enough to guarantee that $\lceil (C_1/D_3)^{2/d} \rceil \sigma^{(1)}(1/4, 1/2) \le 1$, where $\sigma^{(1)}(\lambda, \delta)$ is the constant introduced before (5.40), we can now use the bound (5.40). Combined with (7.4) we get $Pr_p\{W_{\Lambda_r}^{(1)} \ge C_1 s(r)\} \ge 1/2$. Using (7.3), we therefore conclude that there exists a constant $C_3 > 0$ and a $w_0 \in \Lambda_r$ such that

$$Pr_p\{|\mathcal{C}_{\Lambda_r}(w_0)| \ge C_1 s(r)\} \ge \tfrac{1}{2} C_2 \pi_r(p) \ge C_3 \pi_r(p_c), \tag{7.5}$$

where we used Postulate (III) in the last step. Now for any $v \in \Lambda_r$, Λ_r shifted by $v - w_0$ is contained in $\Lambda_{3r} \subset \Lambda$. Therefore for all $v \in \Lambda_r = \Lambda_{\lfloor L_0(p)/4 \rfloor}$ and sufficiently small C_4

$$Pr_p\{|\mathcal{C}_\Lambda(v)| \ge C_4 s(L)\} \ge Pr_p\{|\mathcal{C}_{\Lambda_r}(w_0)| \ge C_1 s(r)\} \ge C_3 \pi_r(p_c) \ge C_2 \pi_L(p_c). \tag{7.6}$$

This proves (7.2) for $\mathbf{j} = \mathbf{0}$ and $x = C_4 \wedge 1$. But then it clearly holds for all $\mathbf{j}$ by translation and for all $0 < x \le C_4 \wedge 1$.

Step ii Now fix k and for brevity write $M = k^d$. Let C_2 and C_4 be such that (7.6) holds. Also fix $x = C_4 \wedge 1 \wedge x_0$ and take $D_7 = D_7(x)$. It is useful to indicate the choice of the $v(\mathbf{j})$ more explicitly in our notation. With some abuse of notation we denote the possible values of $\mathbf{j}$ by $\mathbf{1}, \ldots, \mathbf{M}$, and we occasionaly write $G_k(v(\mathbf{1}), \ldots, v(\mathbf{M}))$ instead of G_k, and similarly for $H_k(v(\mathbf{1}), \ldots, v(\mathbf{M}))$.

We have defined the $\Lambda(\mathbf{j})$ such that they are disjoint. Consequently, for any choice of $v(\mathbf{j})$ in $\widetilde{\Lambda}(\mathbf{j})$, we have by (7.2)

$$Pr_p\{G_k(v(\mathbf{1}),\dots,v(\mathbf{M}))\} = \prod_{\mathbf{j}} Pr_p\{G(\mathbf{j})\} \geq [C_2\pi_L(p_c)]^M,$$

and then by Postulate (VII alt)

$$Pr_p\{G_k(v(\mathbf{1}),\dots,v(\mathbf{M})) \cap H_k(v(\mathbf{1}),\dots,v(\mathbf{M}))\} \geq [D_7C_2\pi_L(p_c)]^M. \tag{7.7}$$

We sum this over all $v(\mathbf{j}) \in \widetilde{\Lambda}(\mathbf{j})$ for $\mathbf{j} \neq \mathbf{1}$. We indicate this sum by $\sum^{(k)}$. We therefore have for some constants C_5, C_6

$$\begin{aligned}\sum{}^{(k)} Pr_p\{G_k(v(\mathbf{1}),\dots,v(\mathbf{M})) \cap H_k(v(\mathbf{1}),\dots,v(\mathbf{M}))\\ \geq [D_7C_2\pi_L(p_c)]^M[2\lfloor L_0(p)/4\rfloor]^{M-1} \geq C_5\pi_L(p_c)[C_6s(L)]^{M-1}.\end{aligned} \tag{7.8}$$

Step iii We next work on an upper bound for the left hand side of (7.8). To this end we note that on the event $G_k \cap H_k$, $v(\mathbf{j})$ is connected to $v(\mathbf{1})$ and therefore to $\partial\Lambda(\mathbf{j})$ whenever $\mathbf{j} \neq \mathbf{1}$. We therefore define

$$\widetilde{V}(\mathbf{j}) = \text{number of } v \in \widetilde{\Lambda}(\mathbf{j}) \text{ which are connected to } \partial\Lambda(\mathbf{j}).$$

We further define

$$I_k = I_k(v(\mathbf{1})) = I[|\mathcal{C}_\Xi(v(\mathbf{1}))| \geq Mxs(L)]. \tag{7.9}$$

Clearly, on the event $G_k(v(\mathbf{1}),\dots,v(\mathbf{M})) \cap H_k(v(\mathbf{1}),\dots,v(\mathbf{M}))$, it holds that

$$|\mathcal{C}_\Xi(v(\mathbf{1}))| \geq \sum_{\mathbf{j}} |\mathcal{C}_{\Lambda(\mathbf{j})}(v(\mathbf{j}))| \geq Mxs(L) \text{ and } v(\mathbf{j}) \leftrightarrow \partial\Lambda(\mathbf{j}),$$

and therefore

$$\begin{aligned}\sum{}^{(k)} Pr_p\{G_k(v(\mathbf{1}),\dots,v(\mathbf{M})) \cap H_k(v(\mathbf{1}),\dots,v(\mathbf{M}))\}\\ \leq E_p\{I_k(v(\mathbf{1}))I[v(\mathbf{1}) \leftrightarrow \partial\Lambda(\mathbf{1})] \prod_{\mathbf{j}\neq\mathbf{1}} \widetilde{V}(\mathbf{j})\}.\end{aligned} \tag{7.10}$$

We continue this inequality. For any $\gamma \geq 0$ the right hand side of (7.10) is at most

$$\begin{aligned}&e^{\gamma M}[s(L)]^{M-1}E_p\{I_k(v(\mathbf{1}))\}\\ &\quad + E_p\Big\{I_k(v(\mathbf{1}))I[v(\mathbf{1}) \leftrightarrow \partial\Lambda(\mathbf{1})] \prod_{\mathbf{j}\neq\mathbf{1}} \widetilde{V}(\mathbf{j}); \prod_{\mathbf{j}\neq\mathbf{1}} \widetilde{V}(\mathbf{j}) \geq e^{\gamma M}[s(L)]^{M-1}\Big\}\\ &\leq e^{\gamma M}[s(L)]^{M-1}E_p\{I_k(v(\mathbf{1}))\}\\ &\quad + e^{-\gamma M}[s(L)]^{-M+1}E_p\Big\{I[v(\mathbf{1}) \leftrightarrow \partial\Lambda(\mathbf{1})] \prod_{\mathbf{j}\neq\mathbf{1}} \widetilde{V}^2(\mathbf{j})\Big\}.\end{aligned}$$

Finally we observe that the $\widetilde{V}(\mathbf{j})$ are independent among each other and of $I[v(\mathbf{1}) \leftrightarrow \partial\Lambda(\mathbf{1})]$, because the $\Lambda(\mathbf{j})$ are disjoint. Moreover, for each $\mathbf{j}$,

$$E_p\{\widetilde{V}^2(\mathbf{j})\} \le C_7 s^2(L),$$

by virtue of (4.6). Therefore

$$\begin{aligned} E_p\{I[v(\mathbf{1}) \leftrightarrow \partial\Lambda(\mathbf{1})] \prod_{\mathbf{j}\neq\mathbf{1}} \widetilde{V}^2(\mathbf{j})\} &= Pr_p\{v(\mathbf{1}) \leftrightarrow \partial\Lambda(\mathbf{1})\} \prod_{\mathbf{j}\neq\mathbf{1}} E_p\{\widetilde{V}^2(\mathbf{j})\} \\ &\le \pi_{\lfloor L/2 \rfloor}(p_c)[C_7 s(L)]^{2M-2}. \end{aligned}$$

Combining these estimates gives

$$\begin{aligned} &\sum{}^{(k)} Pr_p\{G_k(v(\mathbf{1}), \dots, v(\mathbf{M})) \cap H_k(v(\mathbf{1}), \dots, v(\mathbf{M}))\} \\ &\qquad \le e^{\gamma M}[s(L)]^{M-1} E_p\{I_k(v(\mathbf{1}))\} + e^{-\gamma M}[s(L)]^{-M+1}\pi_{\lfloor L/2 \rfloor}(p_c)[C_7 s(L)]^{2M-2}. \end{aligned} \tag{7.11}$$

Step iv In this step we complete the deduction of Postulate (VII) from Postulates (I)–(V) and Postulate (VII alt). From (7.8)–(7.11) we obtain (by means of Postulate (IV))

$$\begin{aligned} &e^{\gamma M}[s(L)]^{M-1} E_p\{I_k(v(\mathbf{1}))\} \\ &\qquad \ge C_5 \pi_L(p_c)\big\{[C_6 s(L)]^{M-1} - e^{-\gamma M}[s(L)]^{-M+1}(C_5 D_3)^{-1} 2^{1+1/\rho_1}[C_7 s(L)]^{2M-2}\big\}. \end{aligned}$$

Choosing γ so large that

$$e^{-\gamma} < C_6 C_7^{-2} \wedge \left[\frac{1}{4} C_5 D_3 2^{-1/\rho_1}\right],$$

we find that

$$E_p\{I_k(v(\mathbf{1}))\} \ge \pi_L(p_c) e^{-\gamma M} \frac{1}{2} C_5 C_6^{M-1}. \tag{7.12}$$

Since, by (7.9), the left hand side is no more than

$$P_{\ge Mxs(L)}(p),$$

and, by (4.20), $\pi_L(p_c) \ge C_2^{-1} P_{\ge s(L)}(p_c) \ge C_2^{-1} P_{\ge s(L)}(p)$, we obtain Postulate (VII).

Step v Even though we finished the deduction of Postulate (VII), we point out here that had we summed over $v(\mathbf{1})$ as well, then the derivation given above would have resulted in

$$\begin{aligned} Pr_p\{&W^{(1)}_{\Lambda_{kL}} \ge C_2 M s(L)\} \\ &= Pr_p\{\exists \text{ a cluster in } \bigcup_{\mathbf{j}} \Lambda(\mathbf{j}) \text{ of size } \ge C_2 M s(L)\} \\ &\ge C_9 e^{-\gamma M} C_6^M. \end{aligned} \tag{7.13}$$

This is basically the estimate (5.83) and we can deduce the lower bound in Theorem 3.5 almost immediately from (7.13) without repeating most of its proof from Postulate (VII).

Also (7.13) can be used to derive the desired counterpart to (3.14), namely, for each fixed K and i

$$\limsup P_{p_n}\Big\{\frac{W^{(i)}_{\Lambda_n}}{E_{p_n}\{W^{(i)}_{\Lambda_n}\}} \le K\Big\} < 1, \tag{7.14}$$

when p_n is inside the scaling window, i.e., when (3.5) holds. To see (7.14) for $i = 1$, fix some large K. Then choose k such that for large n

$$KE_{p_n}\{W^{(1)}_{\Lambda_n}\} \le C_2C_1^{-1}\left(\frac{k}{2}\right)^{1/\rho_2} s(n) \text{ and } kL_0(p) > 2n, \tag{7.15}$$

with C_1 as in (4.17). Such a k exists because $E_{p_n}\{W^{(1)}_{\Lambda_n}\}$ and $s(n)$ are of the same order by Theorem 3.1 i) and p_n is inside the scaling window. Finally choose $p'_n \le (p_n \wedge p_c)$ such that

$$n \le kL_0(p'_n) \le 2n.$$

This can be done by virtue of (2.29). Lemma 4.5 then shows that

$$C_2k^ds(L_0(p'_n)) \ge C_2C_1^{-1}\left(\frac{k}{2}\right)^{1/\rho_2} s(n) \ge KE_{p_n}\{W^{(1)}_{\Lambda_n}\} \text{ (see (7.15))}.$$

Finally, then, by (7.13) for $p = p'_n$,

$$\begin{aligned}&P_{p_n}\{W^{(1)}_{\Lambda_n} \ge KE_{p_n}\{W^{(1)}_{\Lambda_n}\}\} \ge P_{p_n}\{W^{(1)}_{\Lambda_n} \ge C_2k^ds(L_0(p'_n))\}\\ &\ge P_{p'_n}\{W^{(1)}_{\Lambda_n} \ge C_2k^ds(L_0(p'_n))\} \ge C_9e^{-\gamma k^d}C_6^{k^d} > 0.\end{aligned}$$

This proves (7.14) for $i = 1$. For general i a little extra work is needed as in the last few lines of the proof of Theorem 3.5. □

7.2. Proof of Theorem 3.7ii.

We briefly indicate how to derive Postulate (VII alt) in dimension 2. We first show that (7.1) holds when x is sufficiently small. In fact, if K_0, y_0 and j_0 are the constants for which (6.35)–(6.37) hold, then this argument works for $x \le [K_0]^{-1}$. With $M(p, j)$ as in (6.36), we have by a Harris ring construction that for suitable constants $C_1, C_2 > 0$ and all $v(\mathbf{j}) \in \widetilde{\Lambda}(\mathbf{j})$,

$$\begin{aligned}&Pr_p\{\exists \text{ cluster } \mathcal{C} \in \Lambda_{L_0(p)} \text{ which contains } v(\mathbf{j}) \text{ and points}\\ &\qquad v, w \text{ with } v_1 \le j_0y_0L_0(p)/2 < (j_0+1)y_0L_0(p)/2 \le w_1\\ &\qquad \text{and with } |\mathcal{C}| \ge K_0^{-1}s(L_0(p))\}\\ &\ge C_1Pr_p\{M(p, j_0)\}Pr_p\{v(\mathbf{j}) \leftrightarrow \partial B_{L_0(p)}(v(\mathbf{j}))\} \ge \frac{C_1y_0}{8(y_0+1)}\pi_{L_0(p)}(p)\\ &\ge C_2\pi_{L_0(p)}(p_c).\end{aligned} \tag{7.16}$$

On the other hand, by definition of $G(\mathbf{j})$,

$$Pr_p\{G(\mathbf{j})\} \leq P_{\geq xs(L_0(p))}(p).$$

Using the bound

$$P_{\geq s}(p) \leq \pi_n(p) + \frac{1}{s}\sum_{m=1}^{n} |\partial B_m| \left[\pi_{\lfloor m/2 \rfloor}(p)\right]^2,$$

proven in [BCKS99], equation (4.20), one easily shows that there is a constant $C_{10} = C_{10}(x) < \infty$ such that

$$P_{\geq xs(L_0(p))}(p) \leq C_{10}\pi_{L_0(p)}(p_c).$$

Since the $G(\mathbf{j})$ for different $\mathbf{j}$ depend on disjoint regions, they are independent, and

$$Pr_p\{G_k\} \leq [P_{\geq xs(L_0(p))}(p)]^{k^2} \leq [C_{10}\pi_{L_0(p)}(p_c)]^{k^2}. \tag{7.17}$$

Finally, denote the event in the left hand side of (7.16) by $K(\mathbf{j})$, that is

$$\begin{aligned} K(\mathbf{j}) = &\exists \text{ cluster } \mathcal{C} \in \Lambda_{L_0(p)} \text{ which contains } v(\mathbf{j}) \text{ and points} \\ &v, w \text{ with } v_1 \leq j_0 y_0 L_0(p)/2 < (j_0+1) y_0 L_0(p)/2 \leq w_1 \\ &\text{and with } |\mathcal{C}| \geq K_0^{-1} s(L_0(p))\}. \end{aligned}$$

Note that $K(\mathbf{j})$ implies $G(\mathbf{j})$ when $x \leq [K_0]^{-1}$. Therefore another Harris ring construction shows that for some constant $C_{11} > 0$,

$$\begin{aligned} Pr_p\{G_k \cap H_k\} &\geq C_{11}^{k^2} Pr_p\{K(\mathbf{j}) \text{ for all } 1 \leq j_i \leq k, i = 1, 2\} \\ &\geq C_{11}^{k^2} [C_2 \pi_{L_0(p)}(p_c)]^{k^2}, \end{aligned}$$

by virtue of (7.16) and the Harris–FKG inequality. Comparing this with (7.17), we see that

$$Pr_p\{G_k \cap H_k\} \geq [C_{11}C_2/C_{10}]^{k^2} Pr_p\{G_k\}.$$

This completes the proof of (7.1) when $x \leq [K_0]^{-1}$.

For our purposes (7.1) for $0 < x \leq [K_0]^{-1}$ is actually good enough, but it is not hard to obtain (7.1) for general $0 < x \leq 1$ now. In fact, we can apply the same argument as above, provided we first prove the following strengthening of (7.16) for some constant $C_{12} > 0$:

$$\begin{aligned} Pr_p\{&\exists \text{ cluster } \mathcal{C} \in \Lambda_{L_0(p)} \text{ which contains } v(\mathbf{j}) \text{ and points} \\ &v, w \text{ with } v_1 \leq j_0 y_0 L_0(p)/2 < (j_0+1) y_0 L_0(p)/2 \leq w_1 \\ &\text{and with } |\mathcal{C}| \geq s(L_0(p))\} \\ \geq\ & C_{12}\pi_{L_0(p)}(p_c). \end{aligned} \tag{7.18}$$

But (7.18) can be derived exactly as (7.16) from an analogue of (6.37) if we start from

$$\begin{aligned}
&Pr_p\{\exists \text{ cluster } \mathcal{C} \in \Lambda_{L_0(p)} \text{ with } \operatorname{diam}(\mathcal{C}) \geq y_0 L_0(p) \text{ and } |\mathcal{C}| \geq s(L_0(p))\} \\
&\geq Pr_p\{W^{(1)}_{L_0(p)} \geq s(L_0(p))\} \\
&\qquad - Pr_p\{\exists \text{ cluster } \mathcal{C} \subset \Lambda_{L_0(p)} \text{ with } \operatorname{diam}(\mathcal{C}) \leq y_0 L_0(p) \text{ but } |\mathcal{C}| \geq s(L_0(p))\} \\
&\geq Pr_p\{W^{(1)}_{L_0(p)} \geq s(L_0(p))\} - C_1 y_0^{-2} \exp[-C_3 y_0^{-1}] \\
&\geq C_{13} > 0, \qquad (7.19)
\end{aligned}$$

which is valid for sufficiently small $y_0 > 0$ and some constant $C_{13} > 0$. Equation (7.19) is the analogue of (6.35) with $[K_0]^{-1}$ replaced by 1. The reason why we can prove this now, but could not take $[K_0]^{-1} = 1$ in (6.35) to begin with, is that we first needed to show that $Pr_p\{W^{(1)}_{L_0(p)} \geq s(L_0(p))\}$ is bounded away from 0. But this is now available to us from (7.14). As we pointed out before (7.14) only needs (7.1) for $0 < x \leq x_0$ for some $x_0 > 0$, and this we just derived. □

Acknowledgment: The authors wish to thank the the Forschungsinstitut of the ETH in Zürich and the Institute for Advanced Study in Princeton for their hospitality and partial support of the research in this paper. The authors are also grateful for partial support from other sources: C.B. was supported by the Commission of the European Union under the grant CHRX-CT93-0411, J.T.C. by NSF grant DMS-9403842, and H.K. by an NSF grant to Cornell University.

References

[ACCFR83] M. Aizenman, J. T. Chayes, L. Chayes, J. Fröhlich and L. Russo, *On a sharp transition from area law to perimeter law in a system of random surfaces*, Commun. Math. Phys. **92** (1983), 19–69.

[Aiz97] M. Aizenman, *On the number of incipient spanning clusters*, Nucl. Phys. B[FS] **485** (1997), 551-582.

[Ald97] D. Aldous, *Brownian excursions, critical random graphs and the multiplicative coalescent*, Ann. Probab. **25** (1997), 812–854.

[Ale96] K. Alexander, private communication (1996).

[AN84] M. Aizenman and C. M. Newman, *Tree graph inequalities and critical behavior in percolation models*, J. Stat. Phys. **36** (1984), 107–143.

[AS92] N. Alon and J. Spencer, *The Probabilistic Method*, Wiley Interscience, New York, 1992.

[BBCK98] B. Bollobás, C. Borgs, J. T. Chayes and J.-H. Kim, unpublished (1998).

[BCKS99] C. Borgs, J. T. Chayes, H. Kesten and J. Spencer, *Uniform boundedness of critical crossing probabilities implies hyperscaling*, Rand. Struc. Alg. **15** (1999), 368–413.

[BC96] C. Borgs and J. T. Chayes, *On the covariance matrix of the Potts model: A random cluster analysis*, J. Stat. Phys. **82** (1996), 1235–1297.

[BI92-1] C. Borgs and J. Imbrie, *Finite-size scaling and surface tension from effective one-dimensional systems*, Commun. Math. Phys. **145** (1992), 235–280.

[BI92-2] C. Borgs and J. Imbrie, *Crossover finite-size scaling at first order transitions*, J. Stat. Phys. **69** (1992), 487–537.

[BK85] J. van den Berg and H. Kesten, *Inequalities with applications to percolation and reliability*, J. Appl. Probab. **22** (1985), 556–569.

[BoK90] C. Borgs and R. Kotecký, *A rigorous theory of finite-size scaling at first-order phase transitions*, J. Stat. Phys. **61** (1990), 79–110.

[Bol84] B. Bollobás, *The evolution of random graphs*, Trans. Am. Math. Soc. **286** (1984), 257–274.

[Bol85] B. Bollobás, *Random Graphs*, Academic Press, London, 1985.

[CC86] J. T. Chayes and L. Chayes, *Percolation and random media*, Random Systems and Gauge Theories, Les Houches, Session XLIII (K. Osterwalder and R. Stora, eds.), Elsevier, Amsterdam, 1986, pp. 1001-1142.

[CC87] J. T. Chayes and L. Chayes, *On the upper critical dimension in Bernoulli Percolation*, Commun. Math. Phys. **113** (1987), 27–48.

[CCD87] J. T. Chayes, L. Chayes and R. Durrett, *Inhomogeneous percolation problems and incipient infinite clusters*, J. Phys. A: Math. Gen. **20** (1987), 1521–1530.

[CCF85] J. T. Chayes, L. Chayes and J. Fröhlich, *The low-temperature behavior of disordered magnets*, Commun. Math. Phys. **100** (1985), 399–437.

[CCFS86] J. T. Chayes, L. Chayes, D. Fisher and T. Spencer, *Finite-size scaling and correlation lengths for disordered systems*, Phys. Rev. Lett. **57** (1986), 2999-3002.

[CCGKS89] J. T. Chayes, L. Chayes, G. R. Grimmett, H. Kesten and R. H. Schonmann, *The correlation length for the high-density phase of Bernoulli percolation*, Ann. Probab. **17** (1989), 1277–1302.

[Cha98] J. T. Chayes, *Finite-size scaling in percolation*, Doc. Math. J. DMV, Extra Volume ICM III (1998), 113–122.

[CPS96] J. T. Chayes, A. L. Puha and T. Sweet, *Independent and dependent percolation*, IAS-Park City Mathematics Series, Vol. 6: Probability Theory and Applications (Princeton, NJ, 1996), AMS, Providence, RI, 1999, pp. 49–166.

[Con85] A. Coniglio, *Shapes, surfaces and interfaces in percolation clusters*, in Proc. Les Houches Conference on Physics of Finely Divided Matter (M. Daoud and N. Boccara, eds.), Springer-Verlag, 1985, pp. 84–109.

[ER59] P. Erdős and A. Rényi, *On random graphs I*, Publ. Math. (Debrecen) **6** (1959), 290-297.

[ER60] P. Erdős and A. Rényi, *On the evolution of random graphs*, Magy. Tud. Akad. Mat. Kut. Intéz. Közl. **5** (1960), 17–61.

[GM90] G. Grimmett and J. M. Marstrand, *The supercritical phase of percolation is well behaved*, Proc. Roy. Soc. London Ser. A **430** (1990), 439–457.

[Gri99] G. Grimmett, *Percolation*, 2nd edition, Springer-Verlag, New York, 1999.

[Ham57] Hammersley, *Percolation processes. Lower bounds for the critical probability*, Ann. Math. Statist **28** (1957), 790-795.

[Har01] T. Hara, *Critical two-point functions for nearest-neighbour high-dimensional self-avoiding walk and percolation*, preprint in preparation.

[HHS01] T. Hara, R. van der Hofstad and G. Slade, *Critical two-point functions and the lace expansion for spread-out high-dimensional percolation and related models*, preprint (2001).

[Har60] T. E. Harris, *A lower bound for the critical probability in a certain percolation process*, Proc. Cambridge Philos. Soc. **56** (1960), 13–20.

[Jár00] A. Járai, *Incipient infinite percolation clusters in 2D*, preprint (2000).

[JKŁP93] S. Janson, D. E. Knuth, T. Łuczak and B. Pittel, *The birth of the giant component*, Random Struc. Alg. **4** (1993), 233–358.

[Kes82] H. Kesten, *Percolation Theory for Mathematicians*, Birkhäuser, Boston, 1982.

[Kes86] H. Kesten, *The incipient infinite cluster in two-dimensional percolation*, Probab. Theory Rel. Fields **73** (1986), 369–394.

[Kes87] H. Kesten, *Scaling relations for 2D-percolation*, Commun. Math. Phys. **109** (1987), 109–156.

[Łuc90] T. Łuczak, *Component behavior near the critical point of the random graph process*, Rand. Struc. Alg. **1** (1990), 287–310.

[Ngu85] B. G. Nguyen, *Correlation lengths for percolation processes*, PhD thesis, UCLA (1985).

[Ngu88] B. G. Nguyen, *Typical cluster size for two-dimensional percolation processes*, J. Stat. Phys. **50** (1988), 715–726.

[Rus78] L. Russo, *A note on percolation*, Z. Wahrsch. verw. Geb. **43** (1978), 39–48.

[SmW78] R. T. Smythe and J. C. Wierman, *First-Passage Percolation on the Square Lattice*, Springer-Verlag, Berlin-Heidelberg, 1978.

[SW78] P. D. Seymour and D. J. A. Welsh, *Percolation probabilities on the square lattice*, Ann. Discrete Math. 3 (1978), 227–245.